Supplementary Information

N-Terminal Speciation in Native Chemical Ligation

Oliver R. Maguire,[a,b] Jiayun Zhu,[a] William D. G. Brittain,[a] Alexander S. Hudson,[a] Steven L. Cobb[a] and AnnMarie C. O'Donoghue*[a]

[a] Department of Chemistry, Durham University, University Science Laboratories, South Road, Durham DH1 3LE, U.K.

[b] Current Address: Radboud University Nijmegen, Institute for Molecules and Materials, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

Table of Contents

S1 General

S1.1 Instrumentation

UV-Vis Spectrophotometry: A Cary 100 UV-Vis spectrophotometer with a temperature regulated cuvette holder and attached heating unit was used to collect all absorbance spectra.

pH Measurement: A MeterLab^{TM} PHM 290 pH-Stat Controller equipped with a radiometer combination electrode filled with saturated KCl solution was used to measure pH.

Chromatography: Chromatography was performed with a Sunfire C18 column on a Perkin Elmer Series 200 HPLC Apparatus with Autosampler, Pump, UV-Vis Detector and Chromatography Interface. Preparative chromatography was performed with a Discovery Bio wide pore C18-5 column from Supelco (5 μm, 25 cm × 10 mm), using a Perkin-Elmer 200 LC pump coupled to a Waters 486 tuneable absorbance detector set at λ = 220 nm.

Mass Spectrometry: Low resolution mass spectrometry was performed on a TQD mass spectrometer. High resolution mass spectrometry was performed on a Thermo-Finnigan LTQ FT mass spectrometer and a Waters LCT Premier XE mass spectrometer.

S1.2 Materials

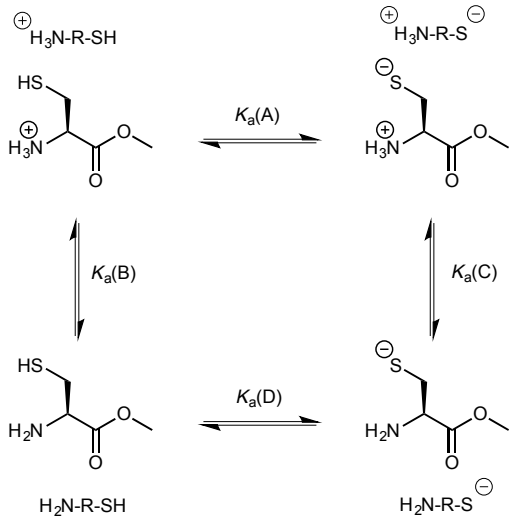
UV-Vis Spectrophotometry: *L*-Cysteine methyl ester, *L*-cysteine methyl ester dimer, *L*-penicillamine methyl ester and H-*L*-Cys-Gly-OH were purchased from Sigma Aldrich. (4*S*)-mercaptoproline methyl ester was prepared as described in the PhD Thesis of Dr Alex Hudson.[S1] Cysteine was purchased from Alfa Aesar. *Tris*(2-carboxyethyl)phosphine (TCEP) was purchased from TCI UK. Volumetric 1 M sodium hydroxide and volumetric hydrochloric acid were purchased from Fisher Scientific UK. All other chemicals used were reagent grade and used without further purification.

Peptide Synthesis: Rink amide resin, DMF, piperidine, *N*,*N*-Diisopropylethylamine, (benzotriazol-1-yloxy)-tripyrrolidinophosphonium hexafluorophosphate (PyBOP), trifluoroacetic acid, triisopropylsilane, Fmoc-Phe-OH, Fmoc-Gly-OH, Fmoc-Cys(Trt)-OH,

Fmoc-Cys(mmt)-OH, Fmoc-Val-OH, Fmoc-Ser(OtBu)-OH, and Fmoc-Pen(Trt)-OH were all provided by the Cobb laboratory.

S2 UV-Vis Spectrophotometric Results

S2.1 Determination of dissociation constants K_a(A) – K_a(D) for cysteine methyl ester (2)

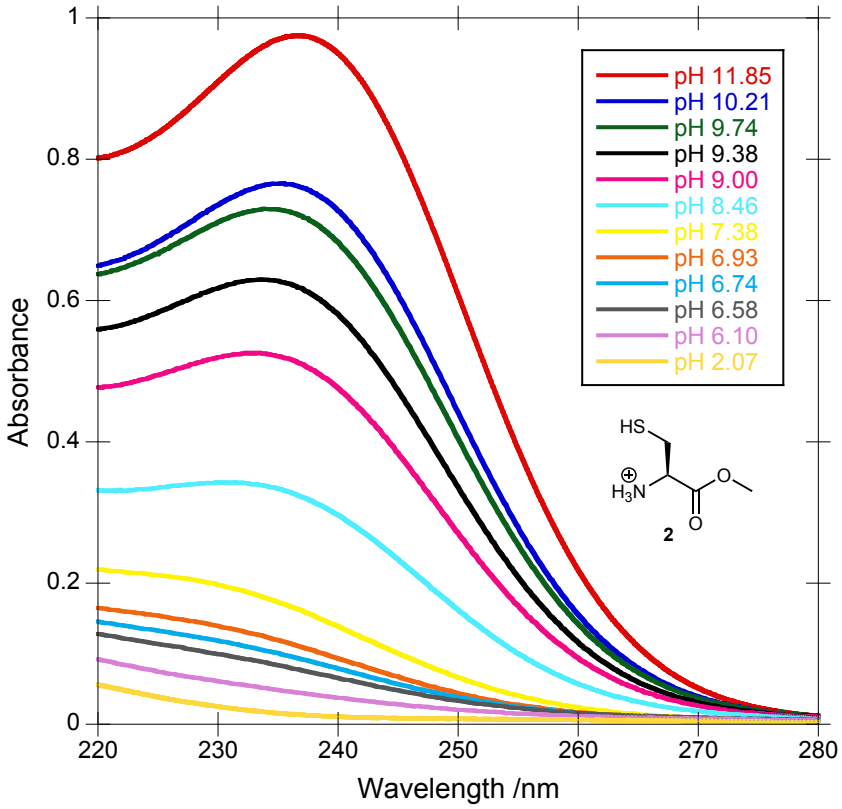


The method employed is a modification of that reported by Benesch for the determination of the dissociation constants for cysteine **1**.[S2] A 0.1 M stock solution of cysteine methyl ester (**2**) was prepared from 85.76 mg (0.500 mmol) of cysteine methyl ester.HCl and 5.00 mL of deionised water. A series of HCl, H_3PO_4, KH_2PO_4, NH_4Cl, $KHCO_3$ and NaOH solutions with varying percentages of free base were prepared and maintained at a constant ionic strength I = 0.3 M (KCl or NaCl). A 0.1 M stock solution of TCEP was prepared using 2.86 mg (9.98×10^{-6} mol) and 100 μL of deionised water. Buffers were chosen with minimal absorption in the vicinity of the λ_{max} = 235 – 237 nm, the n to σ* transition of the thiolate.

Absorbance spectra were measured with a Cary 100 UV-Vis spectrophotometer. The buffer solutions were added to 1.0 mL cuvettes and the cuvettes were placed in a temperature regulated cuvette holder of the UV-Vis spectrophotometer. The solutions were allowed to equilibrate to 25 °C prior to the recording of the absorbance spectrum over the wavelength range 200 – 300 nm (the UV-Vis spectrophotometer had been zeroed against air). 20 μL of 0.1 M TCEP solution was added to the cuvette to give a 2.0 mM solution and the absorption spectrum of the TCEP and buffer solution was recorded. 2 μL of 0.1 M cysteine methyl ester stock solution was added to the cuvette to give a 0.2 mM solution and the absorption spectrum

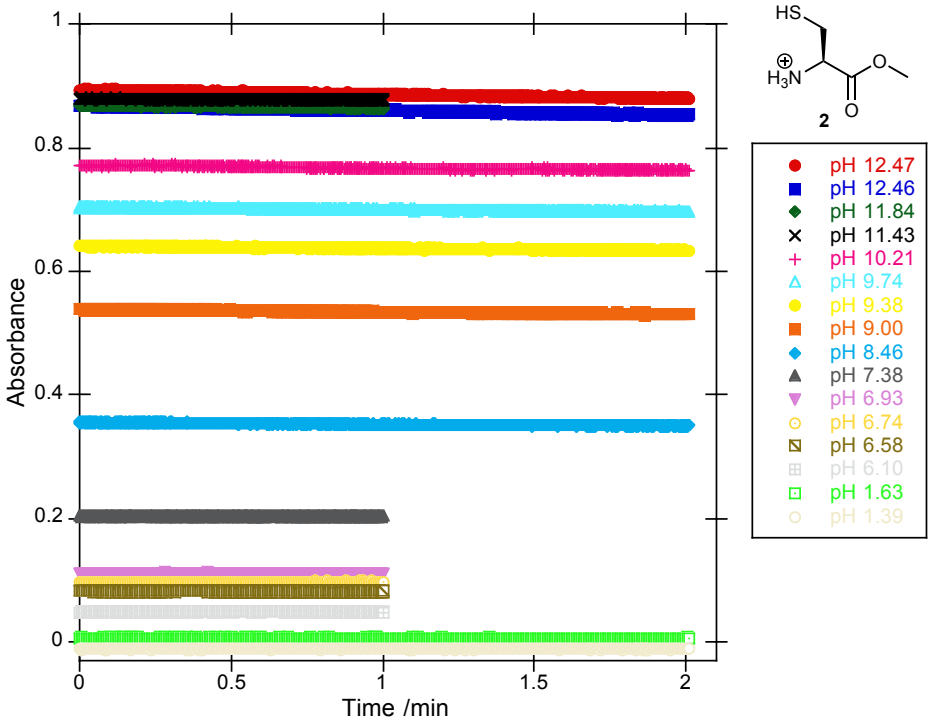
of the cysteine methyl ester was recorded and the absorbance from the buffer solution was manually subtracted. This process was repeated at each pH for every buffer solution (Fig. S1).

Fig. S1: UV-Vis spectra for cysteine methyl ester (**2**) (0.2 mM) in buffer solutions covering the pH range 2.07 – 11.85 at 25 °C and ionic strength I = 0.3 M (NaCl).



The wavelength of maximum absorbance at each pH over the range pH = 2.07 – 11.85 was then identified as λ_{max} = 235 – 237 nm for thiolate and was observed to blue shift with decreasing pH, which is attributed to $^{+}H_3N$-R-S^{-} absorbing at shorter wavelengths than H_2N-R-S^{-}. A single wavelength spectrum was then recorded at the appropriate λ_{max} for the pH of the buffer solution using an identical procedure as above (Fig. S2). The largest value for A_{obs} at the single wavelength recorded was used to determine the acid dissociation constants K_a(A) – K_a(D).

Fig. S2: UV-Vis spectra monitored at a single wavelength (λ = 237 nm) for 0.2 mM cysteine methyl ester (**2**) in buffer solutions covering the pH range 1.39 – 12.47 at 25 °C and ionic strength I = 0.3 M (NaCl).



To determine the dissociation constants K_a(A) – K_a(D), the fraction of total thiol in thiolate form, f_{RS^-}, at each pH is calculated. At pH 12.47, it can be reasonably assumed that the observed absorbance is due to thiolate form only, $(A_{obs})_{max.}$ The f_{RS^-} of thiolate present in solution at lower pHs may then be determined relative to $(A_{obs})_{max.}$ At lower pHs there may also be a small contribution to the absorbance at λ_{max} from the thiol in acid form (R–SH). In order to correct for this, the absorbance value at 237 nm for a solution at pH 1.63, A_{237}^{RSH} (where it is reasonable to assume that all the thiol is in acid form), was subtracted from each observed absorbance in Fig. S2 prior to the determination of f_{RS^-}. The absorbance contribution from R–SH to A_{obs} will decrease as the pH increases from 9.3 % at pH 6.53 to < 0.7 % at pH 11.84. Values for f_{RS^-} as a function of pH were calculated using Eq. 1 (Main text) and are listed in Table S1.

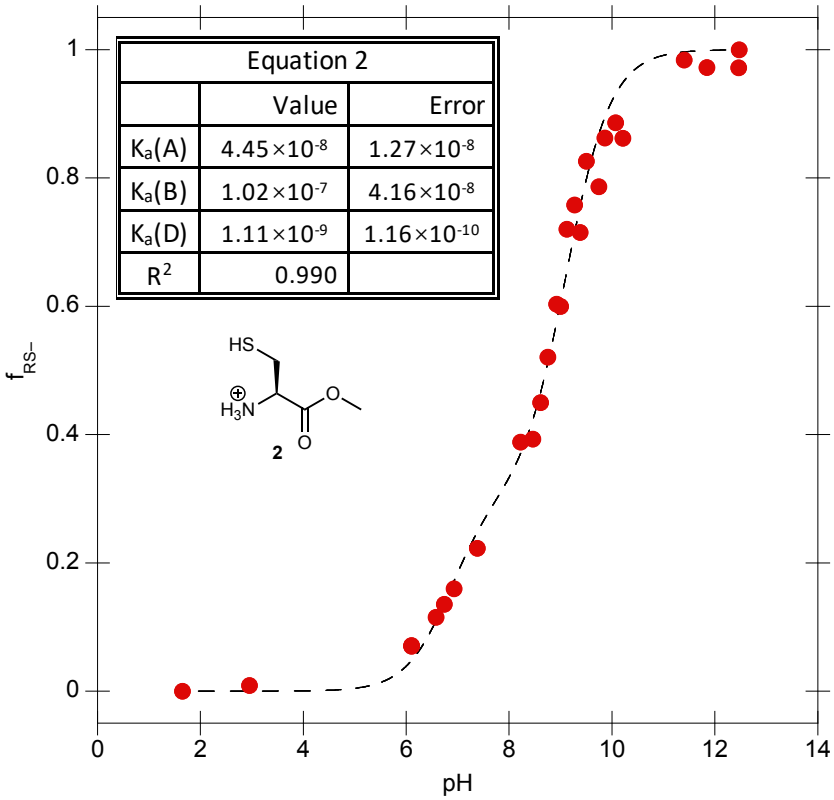
A plot of the f_{RS^-} against pH is shown Fig. S3.

Table S1: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for cysteine methyl ester (**2**) at 25 °C and ionic strength I = 0.3 M (NaCl).

Buffer	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.47	0.893	0.886 [b]	1.00
50 mM NaOH	12.46	0.868	0.861	0.97
30 mM NaOH	11.84	0.869	0.862	0.97
10 mM NaOH	11.40	0.879	0.872	0.98
90 % fb 0.1 M NH_4Cl	10.21	0.771	0.764	0.86
90 % fb 0.1 M $Na_2B_4O_7$	10.07	0.792	0.786	0.89
80 % fb 0.1 M $Na_2B_4O_7$	9.86	0.771	0.764	0.86
70 % fb 0.1 M NH_4Cl	9.74	0.704	0.697	0.79
70 % fb 0.1 M $Na_2B_4O_7$	9.50	0.739	0.732	0.83
50 % fb 0.1 M NH_4Cl	9.38	0.641	0.634	0.72
60 % fb 0.1 M $Na_2B_4O_7$	9.27	0.679	0.672	0.76
50 % fb 0.1 M $Na_2B_4O_7$	9.12	0.645	0.639	0.72
30 % fb 0.1 M NH_4Cl	9.00	0.538	0.532	0.60
40 % fb 0.1 M $Na_2B_4O_7$	8.92	0.542	0.535	0.60
30 % fb 0.1 M $Na_2B_4O_7$	8.75	0.468	0.462	0.52
20 % fb 0.1 M $Na_2B_4O_7$	8.61	0.405	0.399	0.45
10 % fb 0.1 M NH_4Cl	8.46	0.355	0.349	0.39
10 % fb 0.1 M $Na_2B_4O_7$	8.22	0.351	0.344	0.39
80 % fb 0.1 M KH_2PO_4	7.38	0.204	0.197	0.22
60 % fb 0.1 M KH_2PO_4	6.93	0.148	0.142	0.16
50 % fb 0.1 M KH_2PO_4	6.74	0.127	0.120	0.14
40 % fb 0.1 M KH_2PO_4	6.58	0.109	0.102	0.12
20 % fb 0.1 M KH_2PO_4	6.10	0.069	0.062	0.07
20 % fb 0.1 M KH_2PO_4	6.10	0.070	0.063	0.07
90 % fb 0.1 M H_3PO_4	2.95	0.015	0.008	0.01
90 % fb 0.1 M H_3PO_4	2.95	0.015	0.008	0.01
30 mM HCl	1.65	0.007 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S3: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for cysteine methyl ester (**2**) at 25 °C and ionic strength I = 0.3 M (NaCl). K_a(C) = 2.53 × 10^{-9} M. The dotted line is the fit of Eq. 2 to the f_{RS^-} vs. pH data.



Values for the dissociation constants K_a(A), K_a(B) and K_a(D) are estimated from a fitting of data points (●) to Eq. 2 (Main text) (fit shown as dotted line in Fig. S3).

The dissociation constant K_a(C) may then be calculated from the relationship:

$$\frac{K_a(A)}{K_a(B)} = \frac{K_a(D)}{K_a(C)} \qquad \text{(Eq. S1)}$$

Attempts were also made to fit the data to two alternative "Two pK_a + Three Species" models, where the pK_as of the thiol and ammonium are independent of each other (> 2 pK units apart) – see Section S5. However, the fittings were always inferior to the "Four pK_a + Four Species" model used above.

A second repeat of the CysOMe (**2**) experiment was conducted (Fig. S4 + S5, Table S2). The values of dissociation constants, K_a(A) = 4.30 × 10^{-8} M, K_a(B) = 2.16 × 10^{-7} M, K_a(C) = 4.99

$\times 10^{-9}$ M, K_a(D) = 9.94 $\times 10^{-10}$ M were determined for cysteine methyl ester (**2**) by fitting of the data in Fig. S6 to Eq. 2 as described above. K_a(C) was determined using Eq. S1.

A third repeat experiment with CysOMe (**2**) (Fig. S7 + S8, Table S3) had values of K_a(A) = 5.63 $\times 10^{-8}$ M, K_a(B) = 2.77 $\times 10^{-7}$ M, K_a(C) = 5.37 $\times 10^{-9}$ M, K_a(D) = 1.10 $\times 10^{-9}$ M (Fig. S9). All these values are in good agreement with each other.

A fourth repeat experiment with CysOMe (**2**) (Fig. S10 + S11, Table S4) had values of K_a(A) = 2.48 $\times 10^{-8}$ M, K_a(B) = 9.97 $\times 10^{-8}$ M, K_a(C) = 3.50 $\times 10^{-9}$ M, K_a(D) = 8.69 $\times 10^{-10}$ M (Fig. S12). All these values are in good agreement with each other given a multi-parameter fit of this type.

Fig. S4: UV-Vis spectra for 0.2 mM cysteine methyl ester (**2**) in buffer solutions covering the pH range 1.76 – 12.34 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 2)

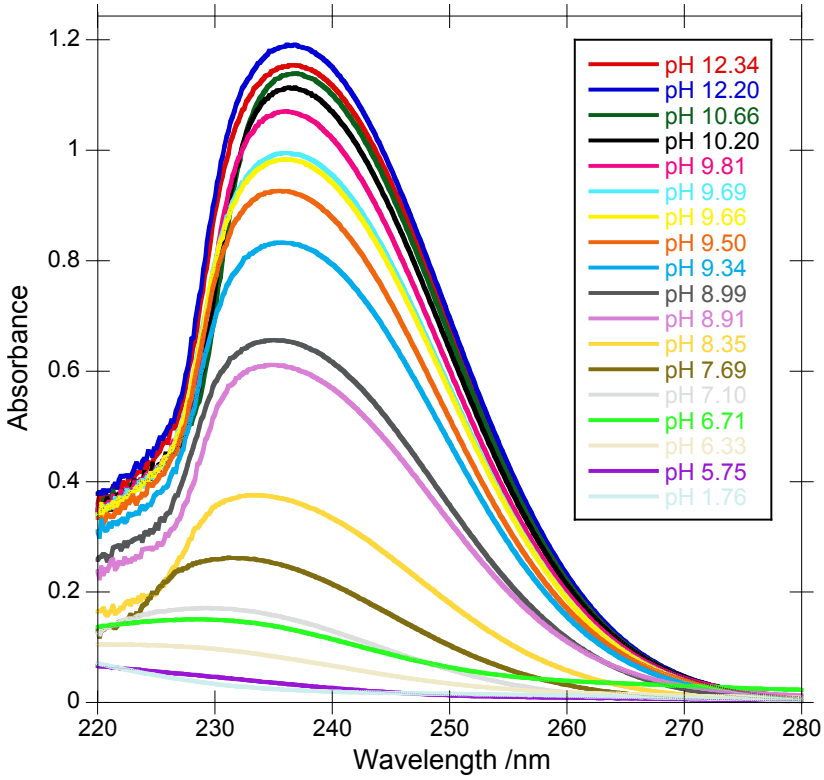


Fig. S5: UV-Vis spectra monitored at a single wavelength (λ = 237 nm) for 0.2 mM cysteine methyl ester (**2**) in buffer solutions covering the pH range 1.49 – 12.34 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 2)

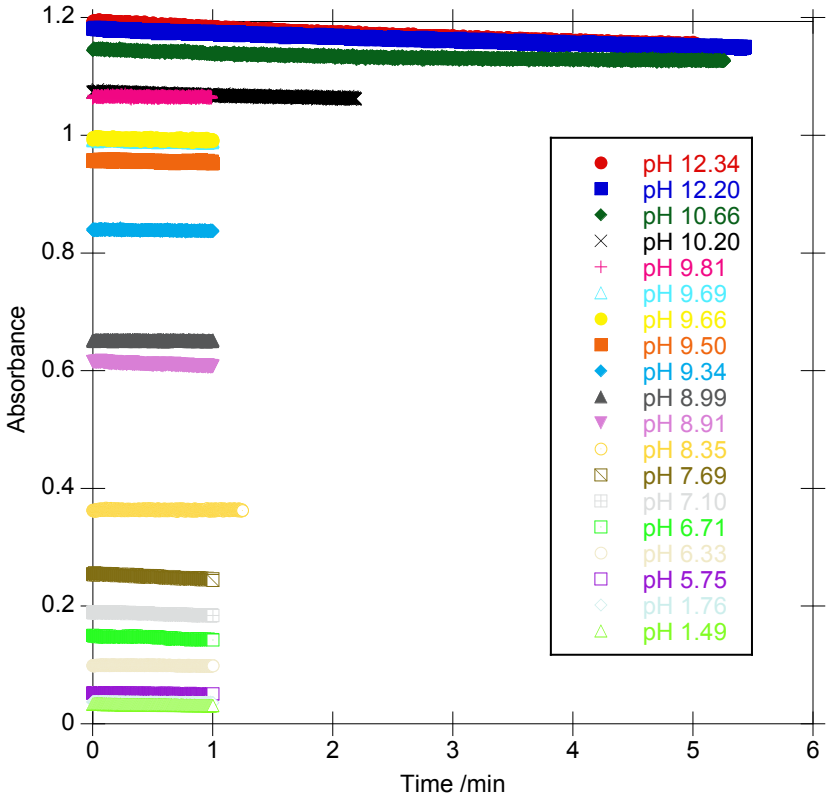


Table S2: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for cysteine methyl ester (**2**) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 2)

Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.34	1.194	1.160 [b]	1.00
30 mM NaOH	12.20	1.181	1.148	0.99
90% fb 0.1 M K_2CO_3	10.66	1.146	1.113	0.96
70% fb 0.1 M K_2CO_3	10.20	1.074	1.040	0.90
50% fb 0.1 M K_2CO_3	9.81	1.066	1.032	0.89
90% fb 0.1 M NH_4Cl	9.69	0.993	0.959	0.83
70% fb 0.1 M NH_4Cl	9.66	0.995	0.961	0.83
30% fb 0.1 M K_2CO_3	9.50	0.958	0.924	0.80
50% fb 0.1 M NH_4Cl	9.34	0.839	0.806	0.70
30% fb 0.1 M NH_4Cl	8.99	0.651	0.618	0.53
10% fb 0.1 M K_2CO_3	8.91	0.615	0.582	0.50
10% fb 0.1 M NH_4Cl	8.35	0.364	0.331	0.29
90% fb 0.1 M KH_2PO_4	7.69	0.255	0.221	0.19
70% fb 0.1 M KH_2PO_4	7.10	0.189	0.156	0.13
50% fb 0.1 M KH_2PO_4	6.71	0.150	0.117	0.10
30% fb 0.1 M KH_2PO_4	6.33	0.100	0.066	0.06
10% fb 0.1 M KH_2PO_4	5.75	0.052	0.019	0.02
30 mM HCl	1.76	0.037	0.004	0.003
50 mM HCl	1.49	0.034 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max}$.

Fig. S6: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for cysteine methyl ester (**2**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 2)

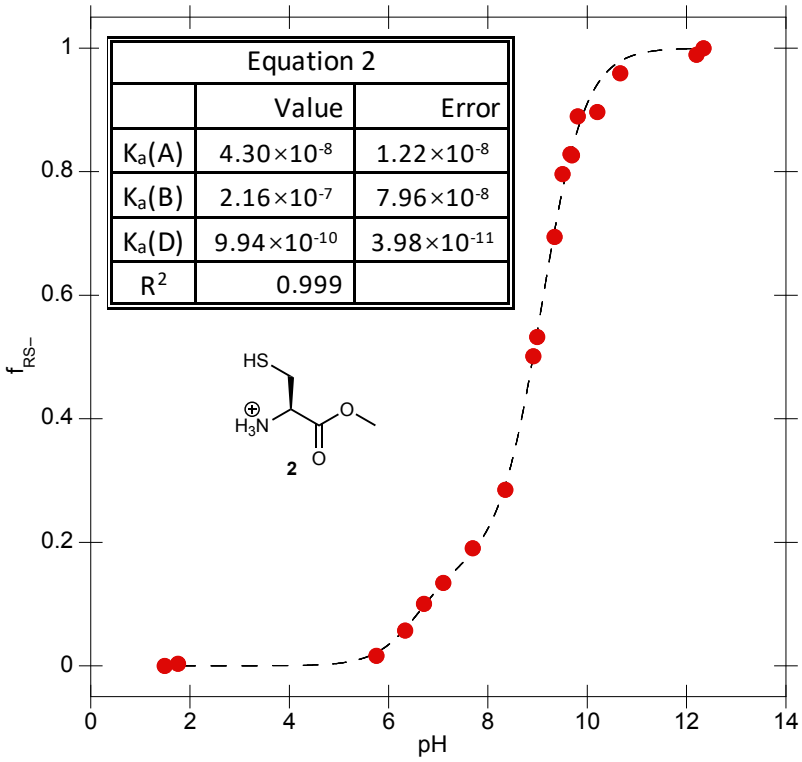


Fig. S7: UV-Vis spectra for 0.2 mM cysteine methyl ester (**2**) in buffer solutions covering the pH range 5.70 – 12.38 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 3)

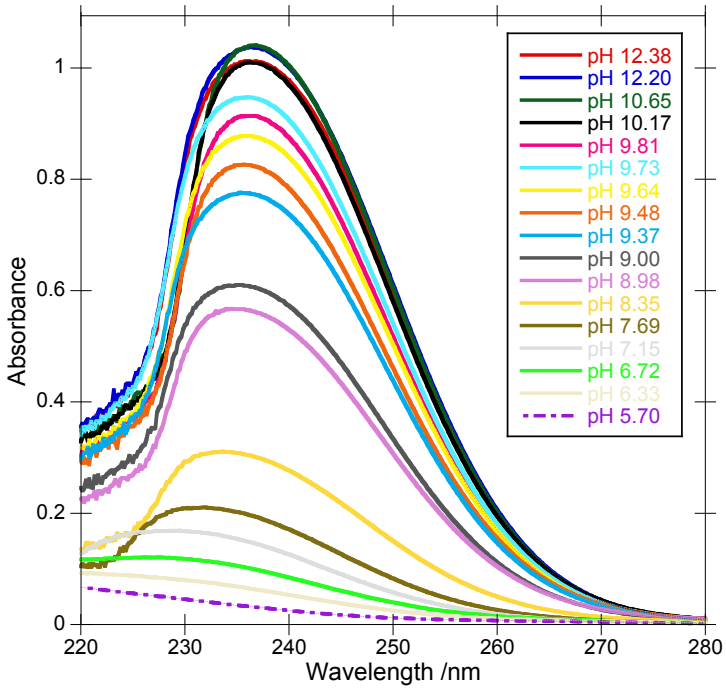


Fig. S8: UV-Vis spectra monitored at a single wavelength (λ = 237 nm) for 0.2 mM cysteine methyl ester (**2**) in buffer solutions covering the pH range 1.39 – 12.38 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 3)

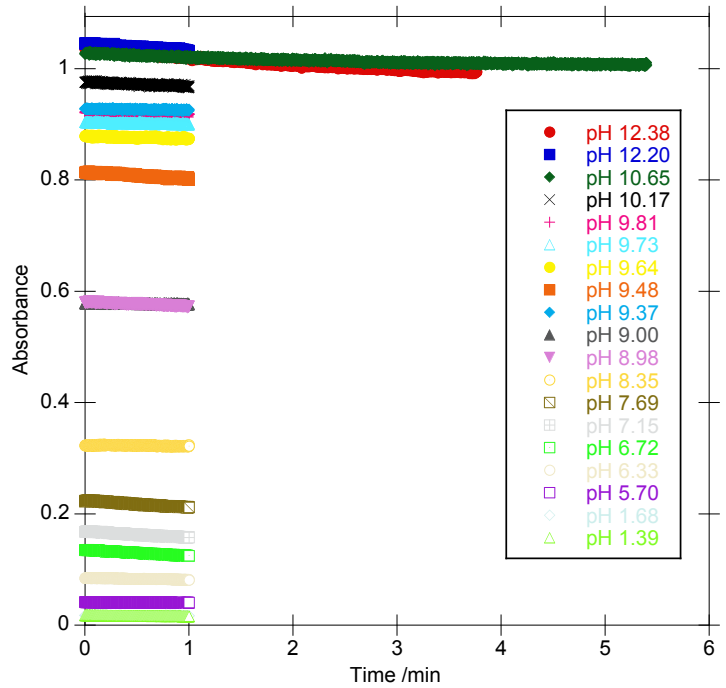


Table S3: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for cysteine methyl ester (**2**) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 3)

Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.38	1.036	1.018 [b]	1.00
30 mM NaOH	12.20	1.045	1.027	1.01
90% fb 0.1 M K_2CO_3	10.65	1.027	1.010	0.99
70% fb 0.1 M K_2CO_3	10.17	0.977	0.959	0.94
50% fb 0.1 M K_2CO_3	9.81	0.922	0.904	0.89
70% fb 0.1 M NH_4Cl	9.73	0.904	0.886	0.87
90% fb 0.1 M NH_4Cl	9.64	0.879	0.861	0.85
30% fb 0.1 M K_2CO_3	9.48	0.813	0.796	0.78
50% fb 0.1 M NH_4Cl	9.37	0.928	0.910	0.89
30% fb 0.1 M NH_4Cl	9.00	0.579	0.561	0.55
10% fb 0.1 M K_2CO_3	8.98	0.580	0.563	0.55
10% fb 0.1 M NH_4Cl	8.35	0.324	0.306	0.30
90% fb 0.1 M KH_2PO_4	7.69	0.224	0.206	0.20
70% fb 0.1 M KH_2PO_4	7.15	0.168	0.150	0.15
50% fb 0.1 M KH_2PO_4	6.72	0.135	0.117	0.12
30% fb 0.1 M KH_2PO_4	6.33	0.085	0.067	0.07
10% fb 0.1 M KH_2PO_4	5.70	0.041	0.024	0.02
30 mM HCl	1.68	0.018	0.000	0.00
50 mM HCl	1.39	0.018 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max}$.

Fig. S9: Repeat experiment for the variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for cysteine methyl ester (**2**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 3)

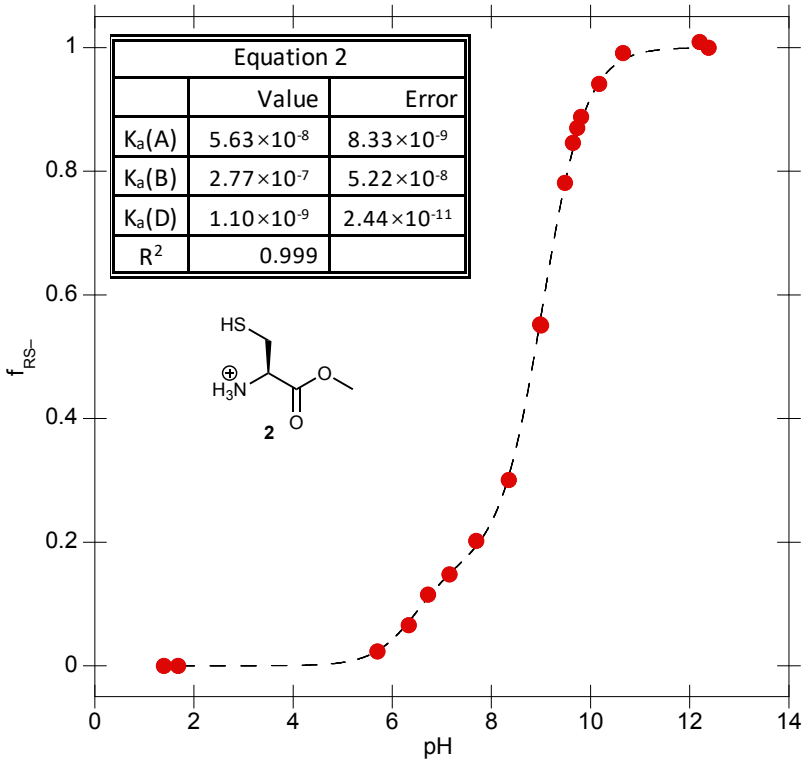


Fig. S10: UV-Vis spectra for 0.2 mM cysteine methyl ester (**2**) in buffer solutions covering the pH range 6.42 – 12.44 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 4)

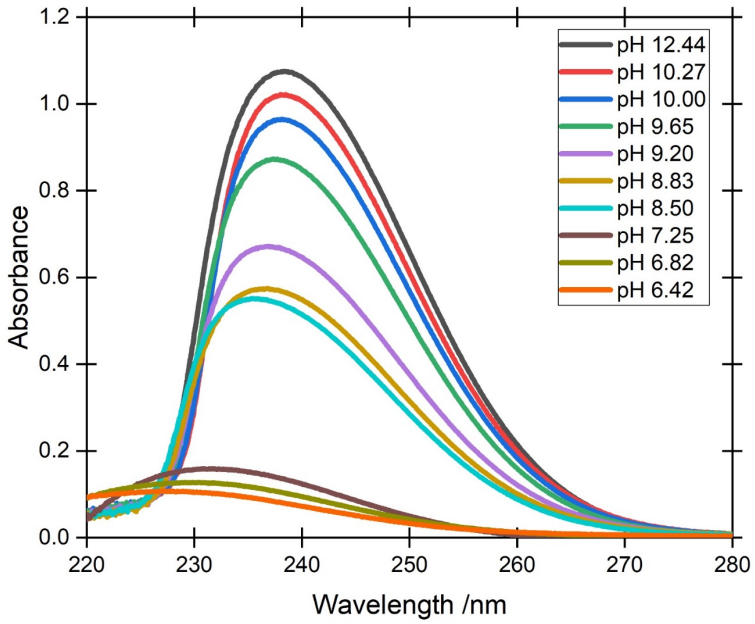


Fig. S11: UV-Vis spectra monitored at a single wavelength (λ = 237 nm) for 0.2 mM cysteine methyl ester (**2**) in buffer solutions covering the pH range 1.38 – 12.44 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 4)

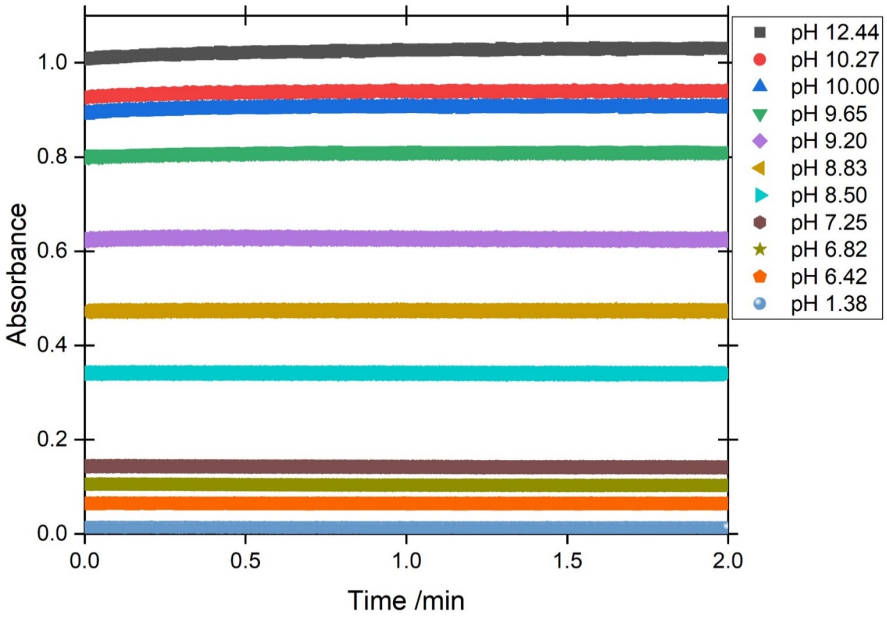
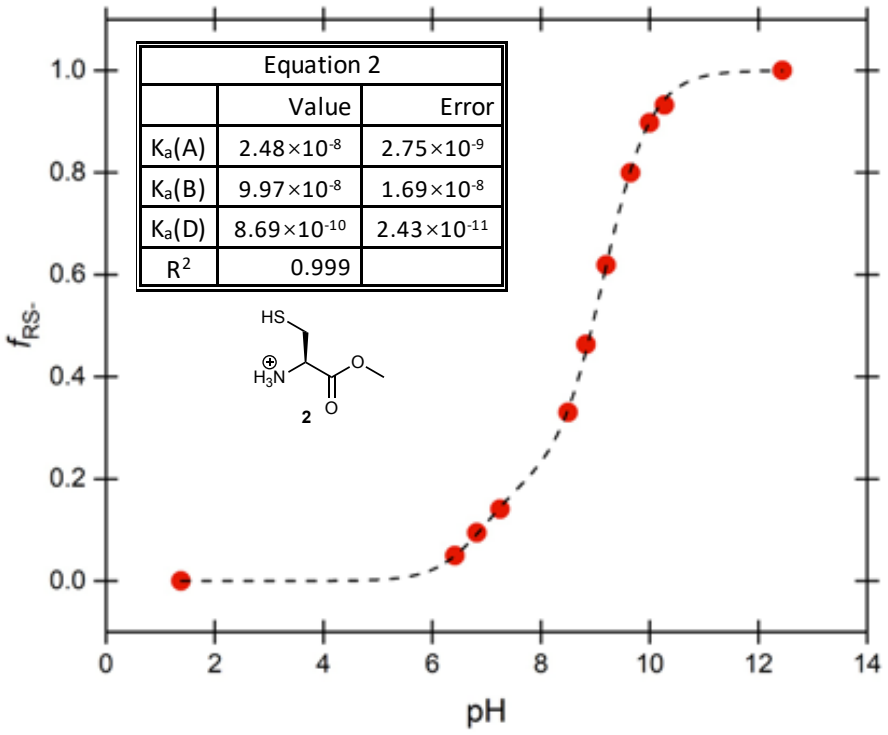


Table S4: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for cysteine methyl ester (**2**) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 4)

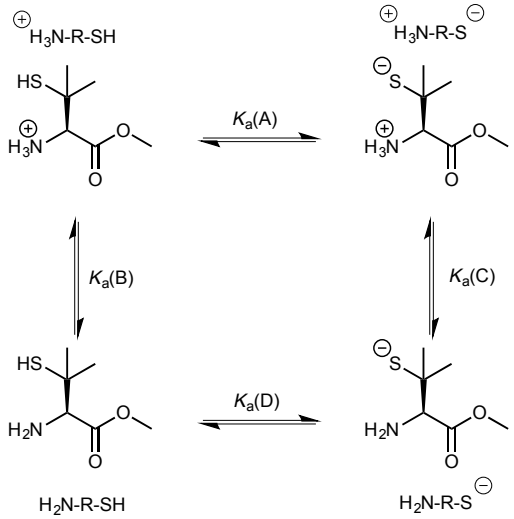
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.44	1.008	0.995[b]	1.00
80% fb 0.1 M K_2CO_3	10.27	0.940	0.927	0.93
60% fb 0.1 M K_2CO_3	10.00	0.906	0.892	0.90
40% fb 0.1 M K_2CO_3	9.65	0.808	0.795	0.80
20% fb 0.1 M K_2CO_3	9.20	0.629	0.616	0.62
30% fb 0.1 M NH_4Cl	8.83	0.475	0.461	0.46
20% fb 0.1 M NH_4Cl	8.50	0.342	0.328	0.33
80% fb 0.1 M K_2HPO_4	7.25	0.153	0.140	0.14
60% fb 0.1 M K_2HPO_4	6.82	0.108	0.094	0.09
40% fb 0.1 M K_2HPO_4	6.42	0.063	0.050	0.05
50 mM HCl	1.38	0.014[a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S12: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for cysteine methyl ester (**2**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat 4)



S2.2 Determination of dissociation constants K_a(A) – K_a(D) for penicillamine methyl ester (3)



A 0.1 M solution of penicillamine methyl ester (**3**) was prepared from 7.99 mg (4.00 x 10^{-5} mol) of penicillamine methyl ester.HCl and 400 µL of deionised water. An identical procedure

to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for penicillamine methyl ester (Fig. S13 + S14, Table S5).

Values of dissociation constants, K_a(A) = 2.15 × 10^{-8} M, K_a(B) = 8.52 × 10^{-8} M, K_a(C) = 1.95 × 10^{-10} M, K_a(D) = 4.86 × 10^{-10} M were determined under non-reducing conditions for penicillamine methyl ester by fitting of the data in Fig. S15 to Eq. 2. K_a(C) was determined using Eq. S1. A repetition experiment under TCEP reducing conditions (Fig. S16 + S17, Table S6) had values of K_a(A) = 4.13 × 10^{-8} M, K_a(B) = 2.44 × 10^{-7} M, K_a(C) = 2.51 × 10^{-9} M, K_a(D) = 4.21 × 10^{-10} M (Fig. S18). Attempt to fit this data to a two pK_a and three species fitting failed to converge upon a solution.

Fig. S13: UV-Vis spectra for 0.2 mM penicillamine methyl ester (3) in buffer solutions covering the pH range 1.42 – 11.82 at 25 °C and ionic strength I = 0.3 M (NaCl).

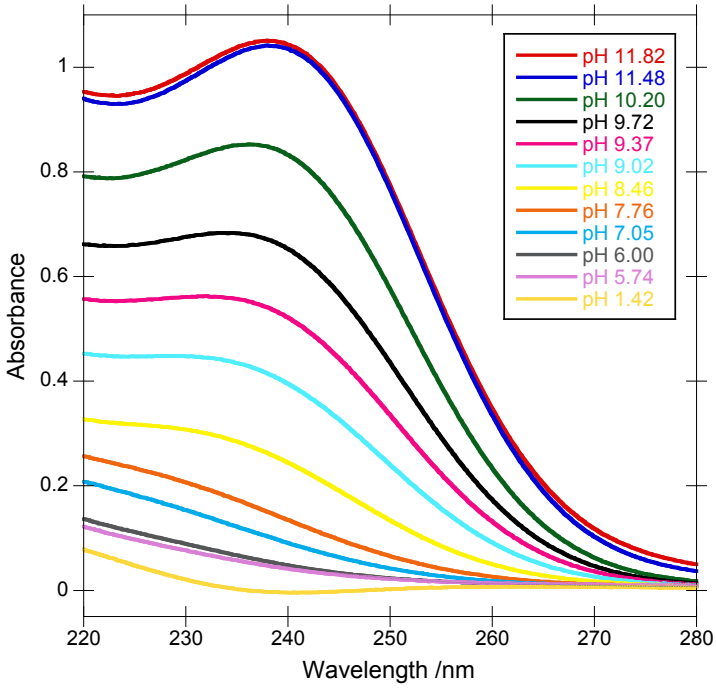


Fig. S14: UV-Vis spectra monitored at a single wavelength (λ = 237 nm) for 0.2 mM penicillamine methyl ester (**3**) in buffer solutions covering the pH range 1.42 – 12.57 at 25 °C and ionic strength I = 0.3 M (NaCl).

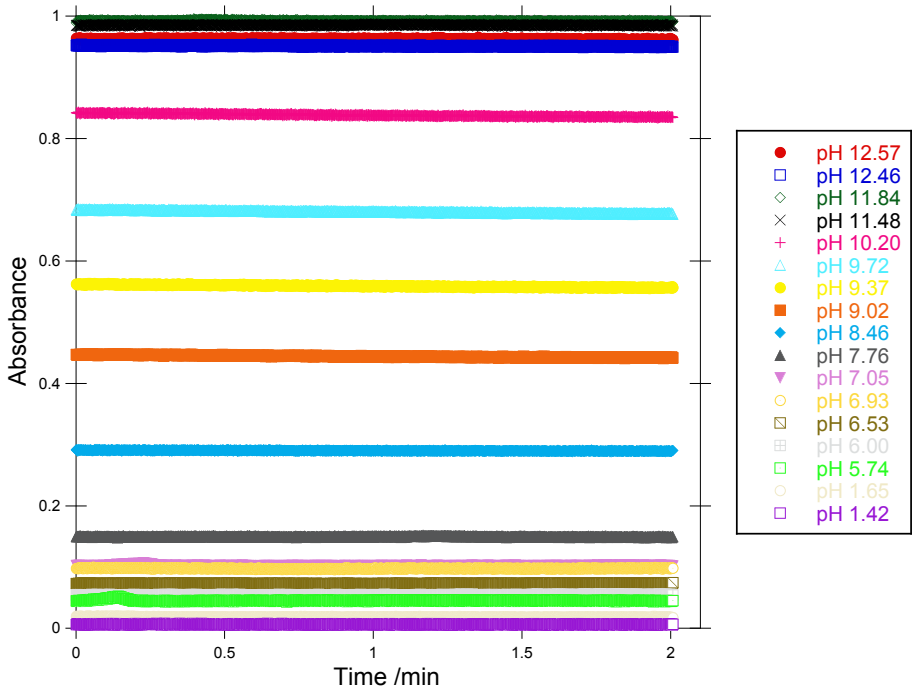


Table S5: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for penicillamine methyl ester (**3**) at 25 °C and ionic strength I = 0.3 M (NaCl).

Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.57	0.963	0.956 [b]	1.00
40 mM NaOH	12.46	0.952	0.945	0.99
30 mM NaOH	11.84	0.992	0.985	1.03
10 mM NaOH	11.49	0.985	0.978	1.02
90% fb 0.1 M NH_4Cl	10.20	0.842	0.835	0.87
70% fb 0.1 M NH_4Cl	9.72	0.684	0.677	0.71
50% fb 0.1 M NH_4Cl	9.37	0.562	0.555	0.58
30% fb 0.1 M NH_4Cl	9.02	0.447	0.440	0.46
10% fb 0.1 M NH_4Cl	8.46	0.291	0.284	0.30
90% fb 0.1 M KH_2PO_4	7.76	0.150	0.143	0.15
70% fb 0.1 M KH_2PO_4	7.05	0.102	0.095	0.10
60% fb 0.1 M KH_2PO_4	6.93	0.098	0.092	0.10
40% fb 0.1 M KH_2PO_4	6.53	0.073	0.066	0.07
20% fb 0.1 M KH_2PO_4	6.00	0.054	0.047	0.05
10% fb 0.1 M KH_2PO_4	5.74	0.045	0.038	0.04
90% fb 0.1 M H_3PO_4	2.97	0.099	0.093	0.10
90% fb 0.1 M H_3PO_4	2.97	0.099	0.092	0.10
70% fb 0.1 M H_3PO_4	2.42	0.099	0.093	0.10
90% fb 0.1 M H_3PO_4	2.42	0.099	0.092	0.10
30 mM HCl	1.42	0.007 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S15: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for penicillamine methyl ester (**3**) (0.2 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).

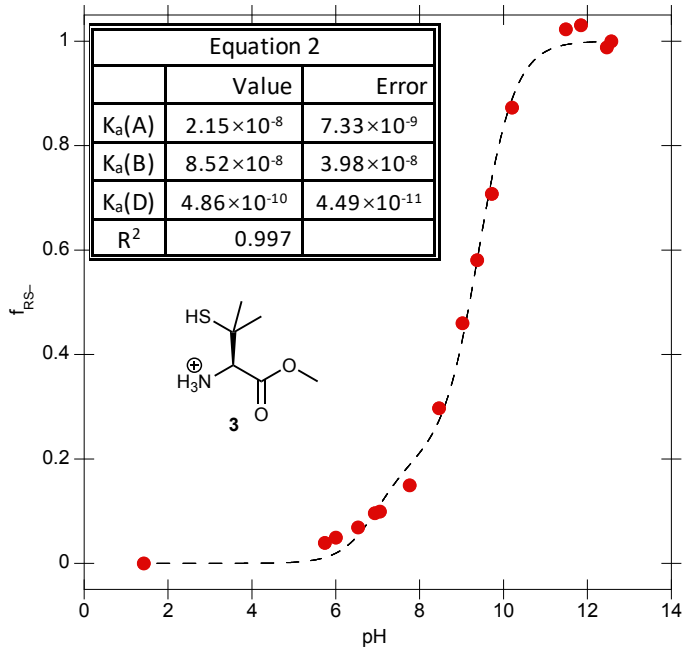


Fig. S16: UV-Vis spectra for 0.2 mM penicillamine methyl ester (**3**) in buffer solutions covering the pH range 1.49 – 12.35 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat)

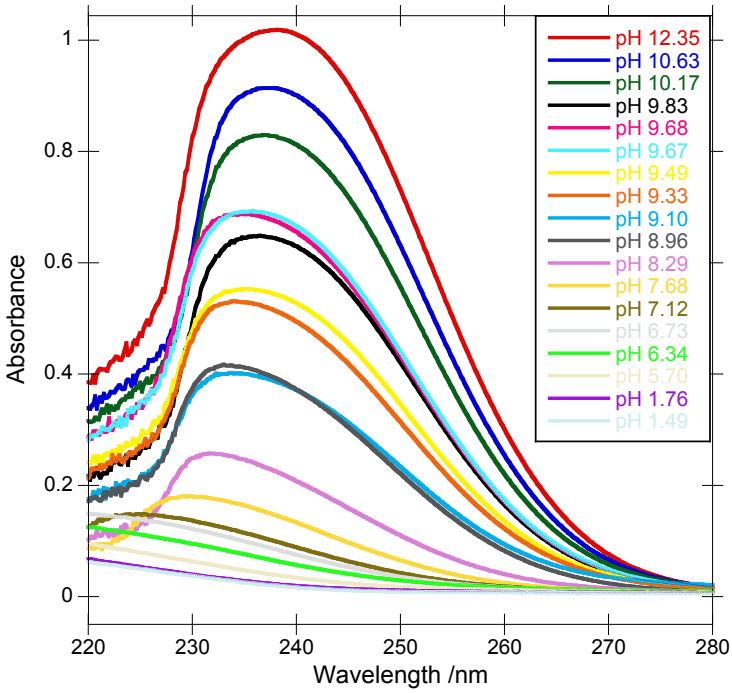


Fig. S17: UV-Vis spectra monitored at a single wavelength (λ = 237 nm) for 0.2 mM penicillamine methyl ester (**3**) in buffer solutions covering the pH range 1.49 – 12.35 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat)

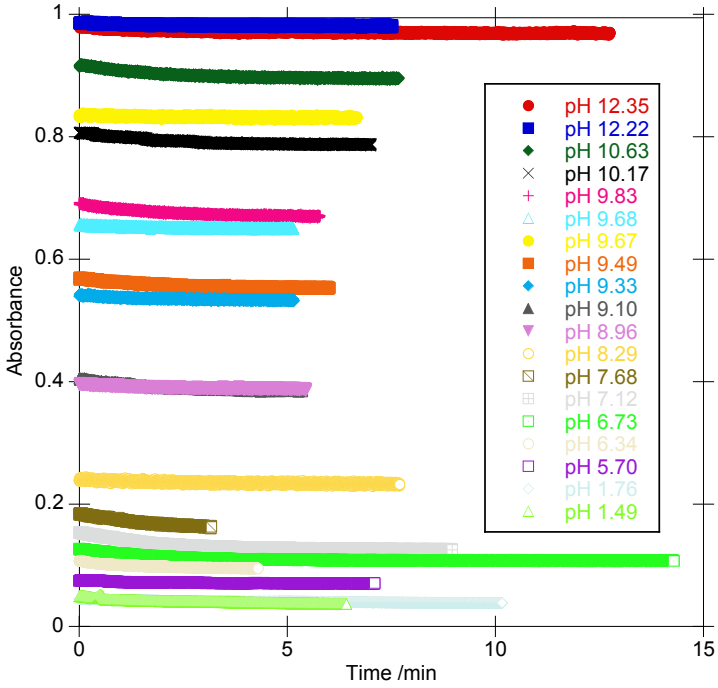
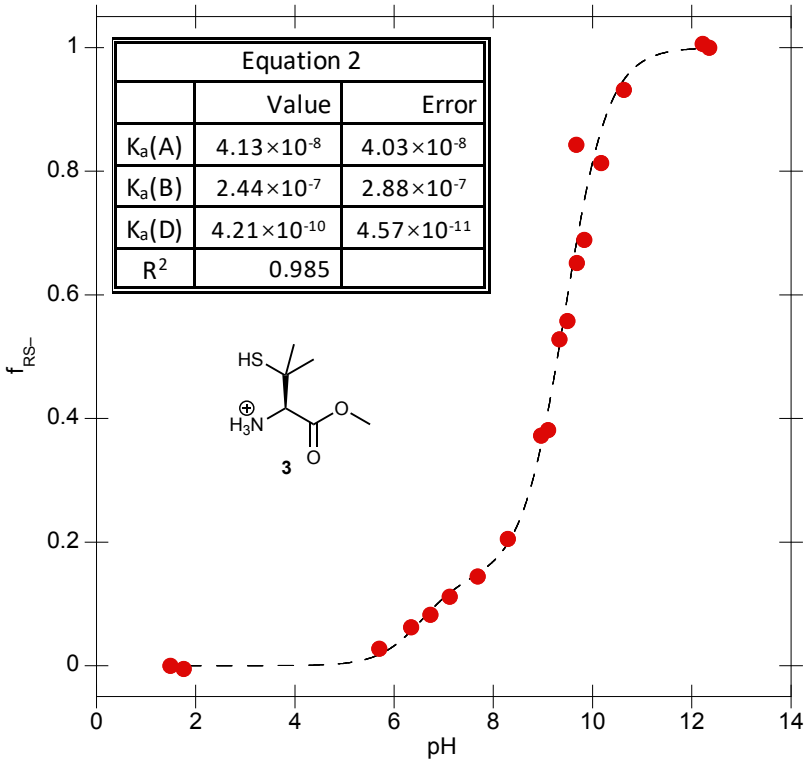


Table S6: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for penicillamine methyl ester (**3**) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat)

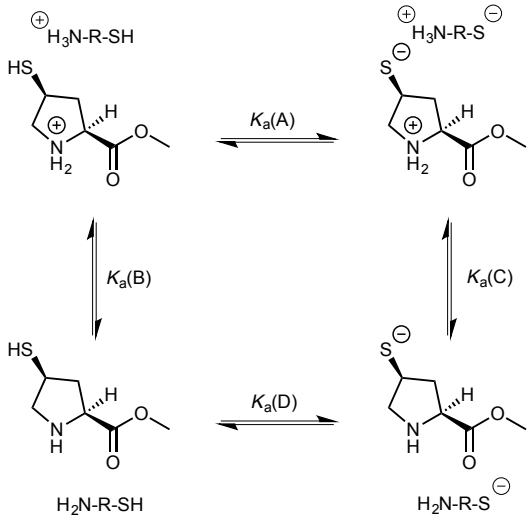
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.35	0.980	0.931 [b]	1.00
30 mM NaOH	12.22	0.986	0.937	1.01
90% fb 0.1 M K_2CO_3	10.63	0.917	0.868	0.93
70% fb 0.1 M K_2CO_3	10.17	0.807	0.757	0.81
50% fb 0.1 M K_2CO_3	9.83	0.691	0.642	0.69
90% fb 0.1 M NH_4Cl	9.68	0.656	0.607	0.65
70% fb 0.1 M NH_4Cl	9.67	0.834	0.785	0.84
30% fb 0.1 M K_2CO_3	9.49	0.569	0.519	0.56
50% fb 0.1 M NH_4Cl	9.33	0.541	0.492	0.53
10% fb 0.1 M K_2CO_3	9.10	0.404	0.355	0.38
30% fb 0.1 M NH_4Cl	8.96	0.396	0.347	0.37
10% fb 0.1 M NH_4Cl	8.29	0.240	0.191	0.21
90% fb 0.1 M KH_2PO_4	7.68	0.184	0.135	0.14
70% fb 0.1 M KH_2PO_4	7.12	0.153	0.104	0.11
50% fb 0.1 M KH_2PO_4	6.73	0.126	0.077	0.08
30% fb 0.1 M KH_2PO_4	6.34	0.107	0.058	0.06
10% fb 0.1 M KH_2PO_4	5.70	0.075	0.026	0.03
30 mM HCl	1.76	0.045	-0.005	-0.01
50 mM HCl	1.49	0.049 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S18: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for penicillamine methyl ester (**3**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl). (Repeat)



S2.3 Determination of dissociation constants K_a(A) – K_a(D) for the (4*S*)-mercaptoproline methyl ester (4)



A 0.1 M solution of (4*S*)-mercaptoproline methyl ester (**4**) was prepared from 7.91 mg (4.00 x 10^{-5} mol) of (4*S*)-mercaptoproline methyl ester. HCl and 400 µL of deionised water. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the

dissociation constants K_a(A) – K_a(D) for (4*S*)-mercaptoproline methyl ester (**4**) (Fig. S19 + S20, Table S7).

Values of dissociation constants, K_a(A) = 7.62 × 10^{-8} M, K_a(B) = 1.28 × 10^{-7} M, K_a(C) = 3.02 × 10^{-9} M, K_a(D) = 1.82 × 10^{-9} M were determined under non-reducing conditions for (4*S*)-mercaptoproline methyl ester by fitting of the data in Fig. S21 to Eq. 2 as described above. K_a(C) was determined using Eq. S1. No visible oxidation was observed during measurements. The oxidation is indicated by a red shift in λ_{max} from 237 nm to 254 – 260 nm.

Fig. S19: UV-Vis spectra for 0.2 mM 4*S*-mercaptoproline methyl ester (**4**) in buffer solutions covering the pH range 1.36 – 12.09 at 25 °C and ionic strength I = 0.3 M (NaCl).

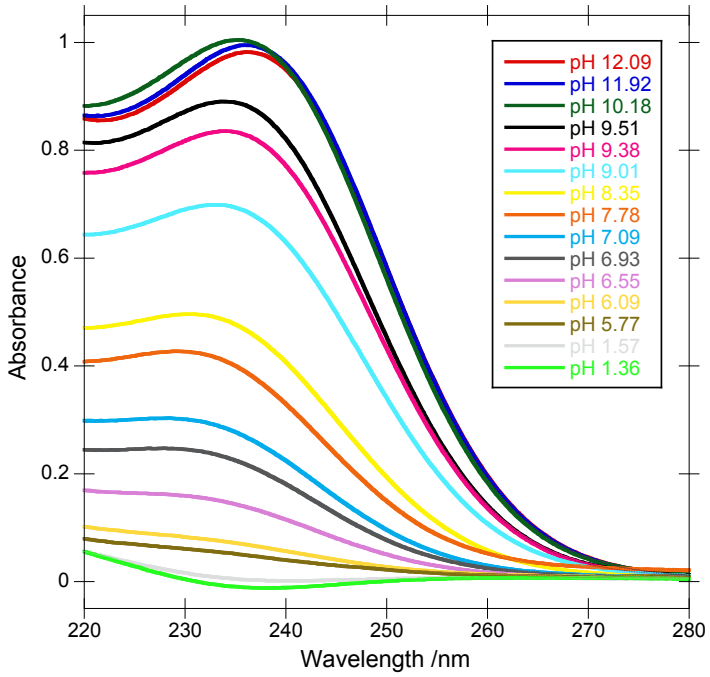


Fig. S20: UV-Vis spectra monitored at a single wavelength (λ = 230 – 236 nm) for 0.2 mM 4*S*-mercaptoproline methyl ester (**4**) in buffer solutions covering the pH range 1.34 – 12.02 at 25 °C and ionic strength I = 0.3 M (NaCl).

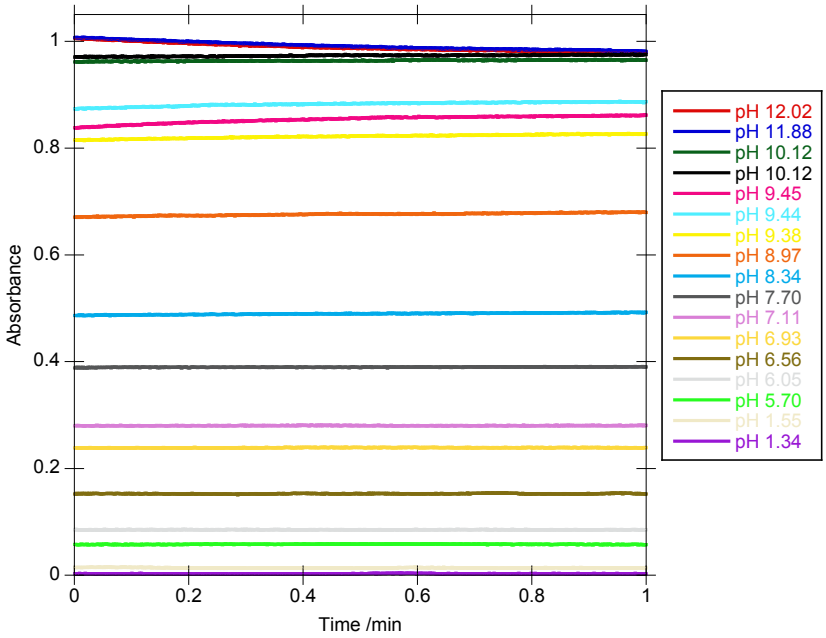
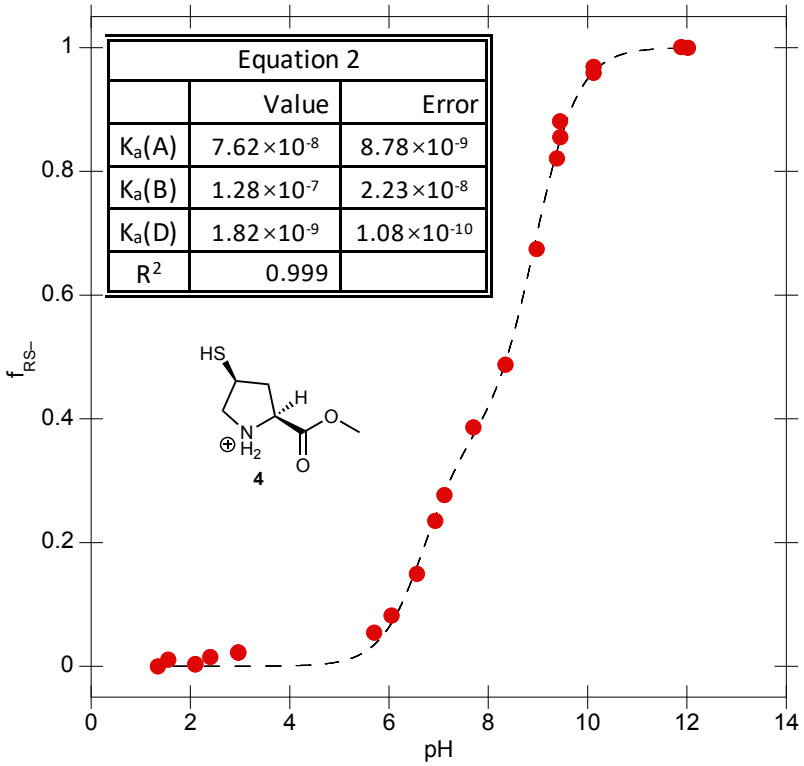


Table S7: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for 4*S*-mercaptoproline methyl ester (**4**) at 25 °C and ionic strength I = 0.3 M (NaCl).

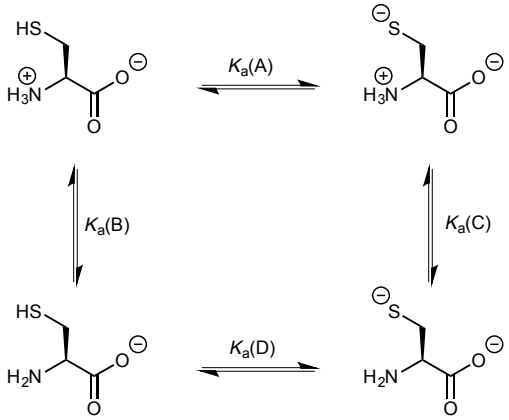
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.02	1.006	1.003 [b]	1.00
30 mM NaOH	11.88	1.007	1.004	1.00
90% fb 0.1 M NH_4Cl	10.12	0.965	0.963	0.96
90% fb 0.1 M NH_4Cl	10.12	0.975	0.972	0.97
70% fb 0.1 M NH_4Cl	9.45	0.861	0.858	0.86
70% fb 0.1 M NH_4Cl	9.44	0.887	0.884	0.88
50% fb 0.1 M NH_4Cl	9.38	0.826	0.824	0.82
30% fb 0.1 M NH_4Cl	8.97	0.680	0.677	0.67
10% fb 0.1 M NH_4Cl	8.34	0.492	0.489	0.49
90% 0.1 M KH_2PO_4	7.70	0.390	0.388	0.39
70% fb 0.1 M KH_2PO_4	7.11	0.280	0.278	0.28
60% fb 0.1 M KH_2PO_4	6.93	0.239	0.236	0.24
40% fb 0.1 M KH_2PO_4	6.56	0.153	0.150	0.15
20% fb 0.1 M KH_2PO_4	6.05	0.085	0.083	0.08
10% fb 0.1 M KH_2PO_4	5.70	0.058	0.055	0.05
90% fb 0.1 M H_3PO_4	2.96	0.025	0.023	0.02
90% fb 0.1 M H_3PO_4	2.96	0.025	0.022	0.02
70% fb 0.1 M H_3PO_4	2.40	0.018	0.015	0.02
70% fb 0.1 M H_3PO_4	2.40	0.018	0.015	0.02
50% fb 0.1 M H_3PO_4	2.10	0.006	0.003	0.00
50% fb 0.1 M H_3PO_4	2.10	0.007	0.004	0.00
30% fb 0.1 M H_3PO_4	1.85	-0.005	-0.007	-0.01
30% fb 0.1 M H_3PO_4	1.85	-0.005	-0.008	-0.01
30 mM HCl	1.55	0.013	0.011	0.01
50 mM HCl	1.34	0.003 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S21: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for (4*S*)-mercaptoproline methyl ester (**4**) (0.2 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S2.4 Determination of dissociation constants K_a(A) – K_a(D) for cysteine (1)



A 0.1 M solution of cysteine (**1**) was prepared from 4.86 mg (4.00 x 10^{-5} mol) of cysteine and 400 µL of deionised water. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for cysteine (Fig. S22 + S23, Table S8).

Values of dissociation constants, K_a(A) = 3.88 × 10^{-9} M, K_a(B) = 3.44 × 10^{-9} M, K_a(C) = 1.32 × 10^{-10} M, K_a(D) = 1.50 × 10^{-10} M were determined under reducing conditions for cysteine by fitting of the data in Fig. S24 to Eq.2 as described above. K_a(C) was determined using Eq. S1.

Fig. S22: UV-Vis spectra for 0.2 mM cysteine (**1**) in buffer solutions covering the pH range 6.69 – 12.39 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

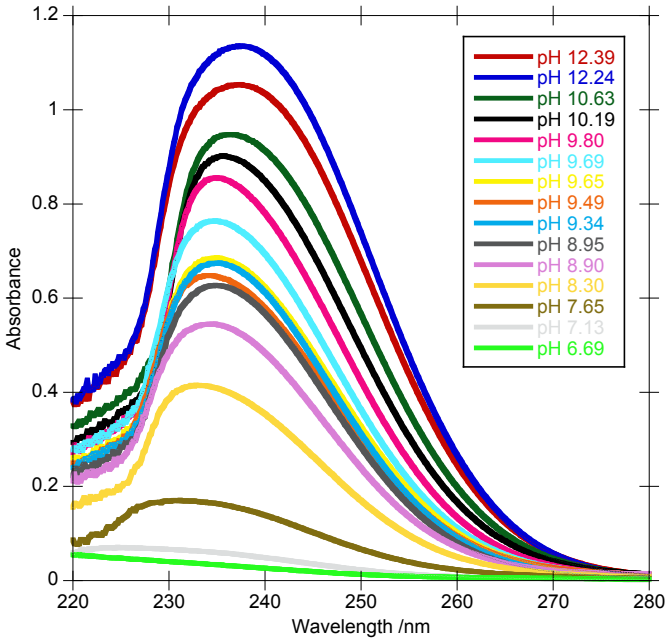
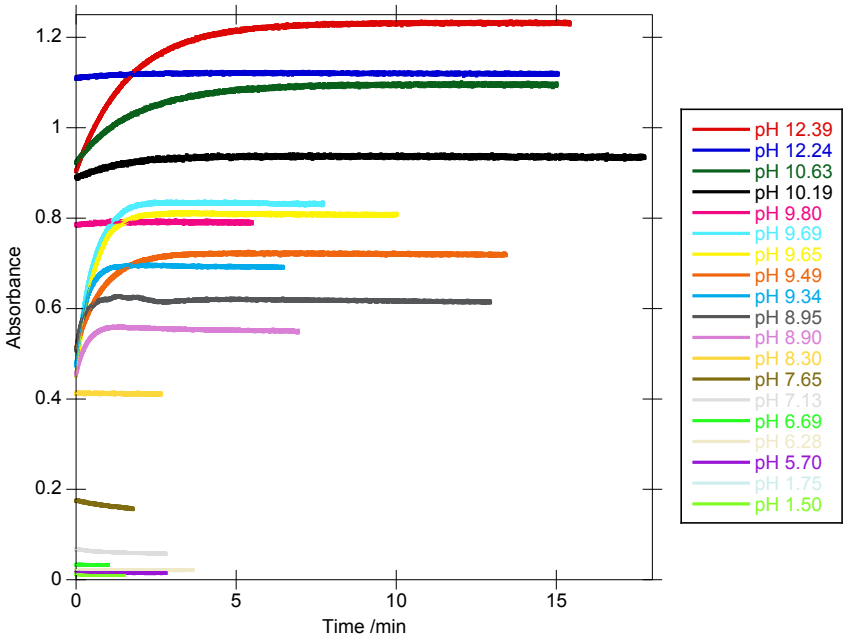


Fig. S23: UV-Vis spectra monitored at a single wavelength (λ = 230 – 236 nm) for 0.2 mM cysteine (**1**) in buffer solutions covering the pH range 1.50 – 12.39 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).



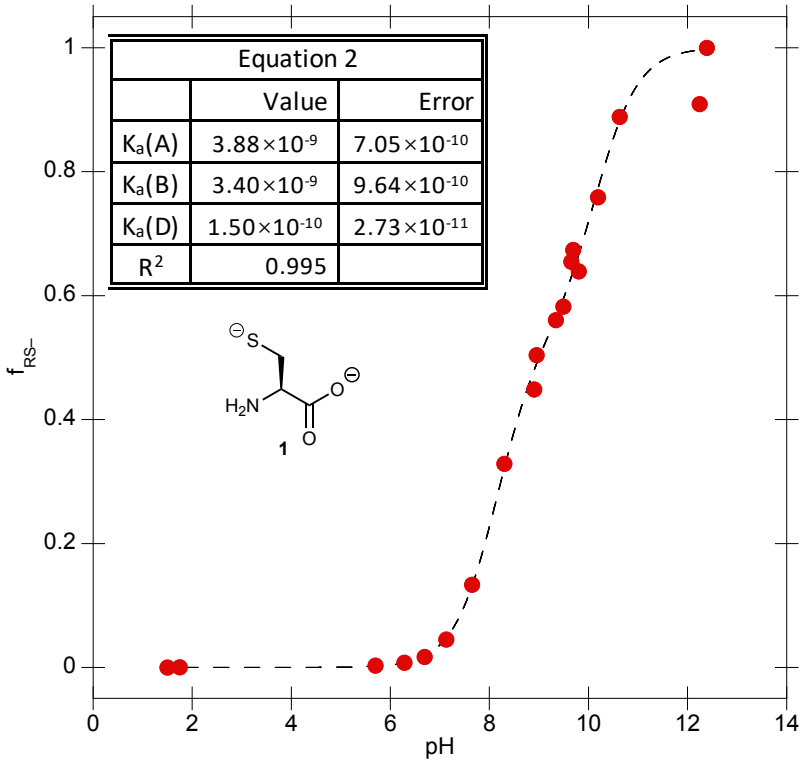
The rise in absorbance over time observed at higher pHs in Fig. S23 indicates that either the cysteine (**1**) had oxidised in the solid form or that the cysteine had oxidised in solution prior to the addition of TCEP solution. To ensure that full reduction of disulfide bonds occurred a fresh TCEP solution was prepared for every measurement from solid TCEP.HCl and the appropriate buffer. The rise in absorbance as the experiment progressed is a result of the TCEP reducing the disulfide bonds. After rising the absorbance then plateaus indicative of full reduction of the disulfide bonds and a consistent thiol(ate) concentration.

Table S8: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for cysteine (**1**) at 25 °C and ionic strength I = 0.3 M (NaCl).

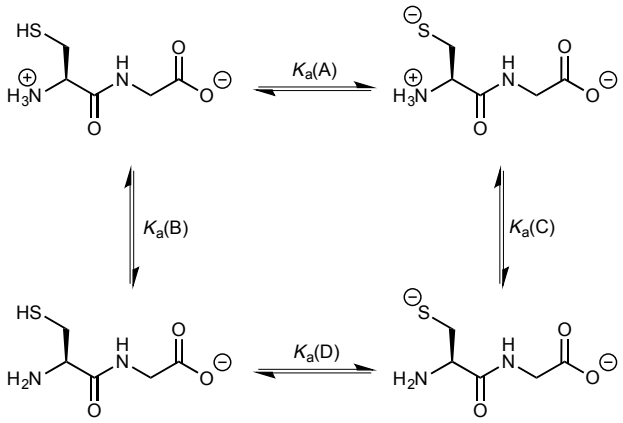
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.39	1.231	1.219 [b]	1.00
30 mM NaOH	12.24	1.121	1.108	0.91
90% fb 0.1 M K_2CO_3	10.63	1.095	1.083	0.89
70% fb 0.1 M K_2CO_3	10.19	0.937	0.925	0.76
50% fb 0.1 M K_2CO_3	9.80	0.792	0.779	0.64
90% fb 0.1 M NH_4Cl	9.69	0.834	0.822	0.67
70% fb 0.1 M NH_4Cl	9.65	0.811	0.799	0.66
30% fb 0.1 M K_2CO_3	9.49	0.722	0.710	0.58
50% fb 0.1 M NH_4Cl	9.34	0.696	0.683	0.56
30% fb 0.1 M NH_4Cl	8.95	0.626	0.614	0.50
10% fb 0.1 M K_2CO_3	8.90	0.559	0.547	0.45
10% fb 0.1 M NH_4Cl	8.30	0.413	0.401	0.33
90% fb 0.1 M KH_2PO_4	7.65	0.175	0.163	0.13
70% fb 0.1 M KH_2PO_4	7.13	0.067	0.055	0.04
50% fb 0.1 M KH_2PO_4	6.69	0.033	0.021	0.02
30% fb 0.1 M KH_2PO_4	6.28	0.022	0.010	0.01
10% fb 0.1 M KH_2PO_4	5.70	0.016	0.004	0.003
30 mM HCl	1.75	0.012	0.000	0.000
50 mM HCl	1.50	0.012 [a]	0.000	0.000

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S24: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for cysteine (**1**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S2.5 Determination of dissociation constants K_a(A) – K_a(D) H-Cys-Gly-OH (5)



A 0.1 M solution of H-Cys-Gly-OH (**5**) was prepared from 7.13 mg (4.00 x 10^{-5} mol) of H-Cys-Gly-OH and 400 µL of deionised water. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for cysteine (Fig. S25 + S26, Table S9).

Values of dissociation constants, K_a(A) = 1.07 × 10^{-8} M, K_a(B) = 9.68 × 10^{-8} M, K_a(C) = 4.57 × 10^{-9} M, K_a(D) = 5.04 × 10^{-10} M were determined under reducing conditions for H-Cys-Gly-OH by fitting of the data in Fig. S27 to Eq.2 as described above. K_a(C) was determined using Eq. S1.

Fig. S25: UV-Vis spectra for 0.2 mM H-Cys-Gly-OH (**5**) in buffer solutions covering the pH range 1.44 – 12.42 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

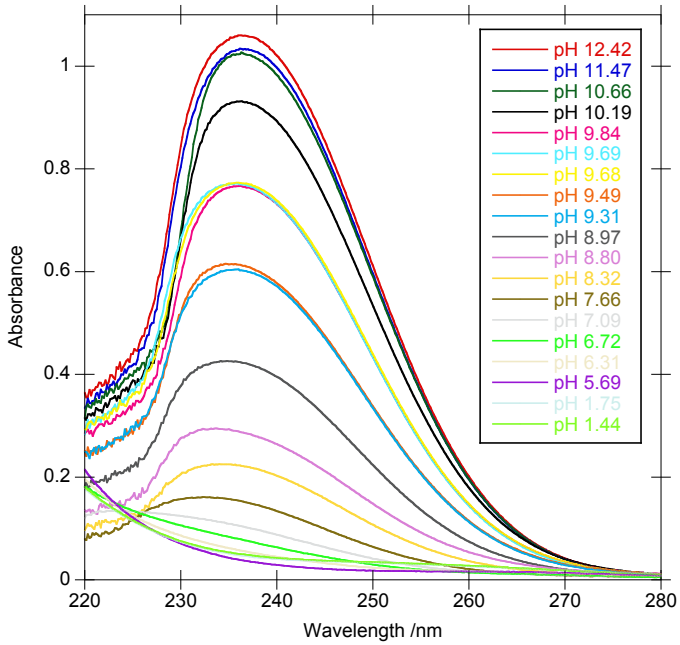
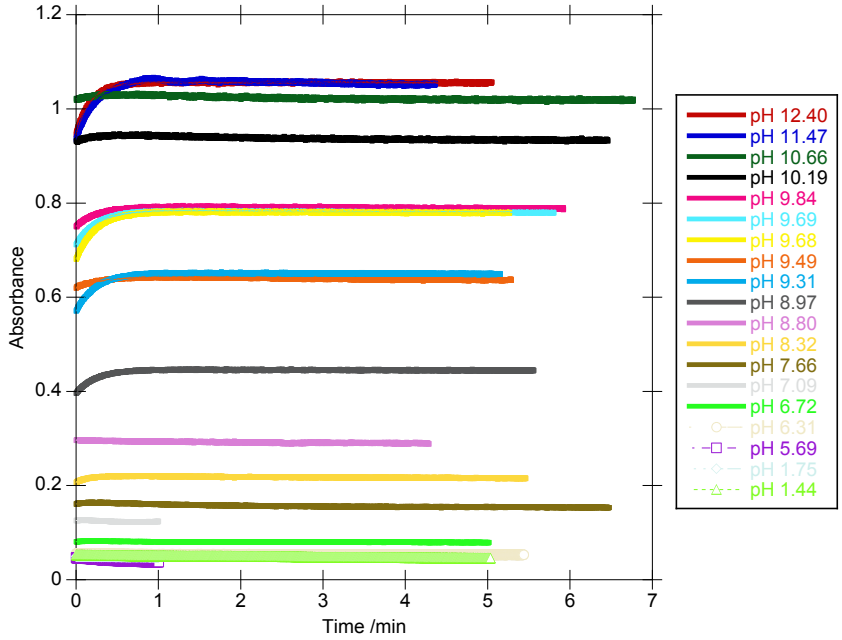


Fig. S26: UV-Vis spectra monitored at a single wavelength (λ = 230 – 236 nm) for 0.2 mM H-Cys-Gly-OH (**5**) in buffer solutions covering the pH range 1.44 – 12.40 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).



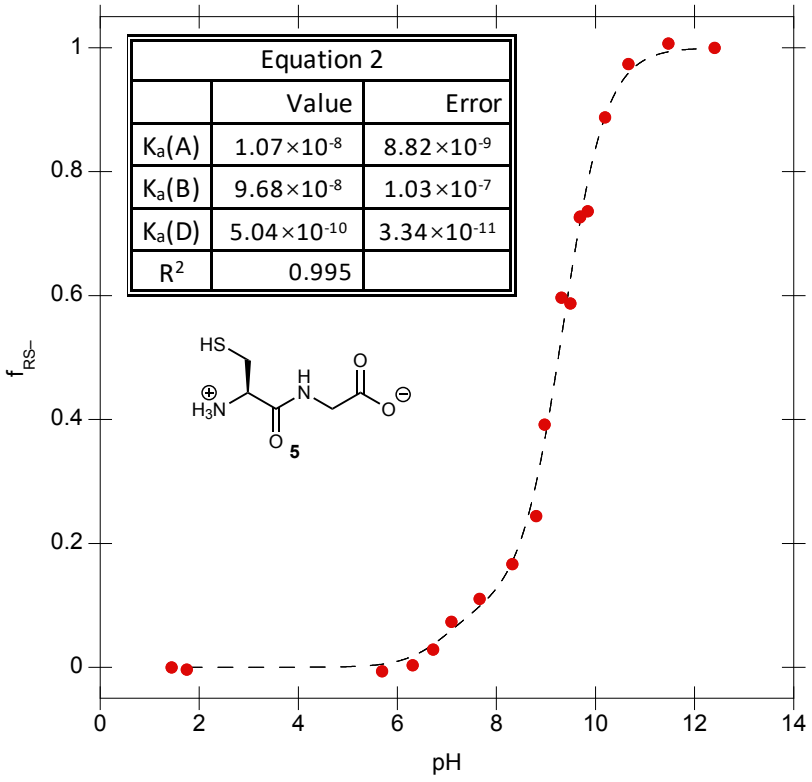
The rise in absorbance over time observed at higher pHs in Fig. S26 indicates that either the H-Cys-Gly-OH (**5**) had oxidised in the solid form or that the H-Cys-Gly-OH had oxidised in solution prior to the addition of TCEP solution. To ensure that full reduction of disulfide bonds occurred a fresh TCEP solution was prepared for every measurement from solid TCEP.HCl and the appropriate buffer. The rise in absorbance as the experiment progressed is a result of the TCEP reducing the disulfide bonds. After rising the absorbance then plateaus indicative of full reduction of the disulfide bonds and a consistent thiol concentration.

Table S9: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for H-Cys-Gly-OH (**5**) at 25 °C and ionic strength I = 0.3 M (NaCl).

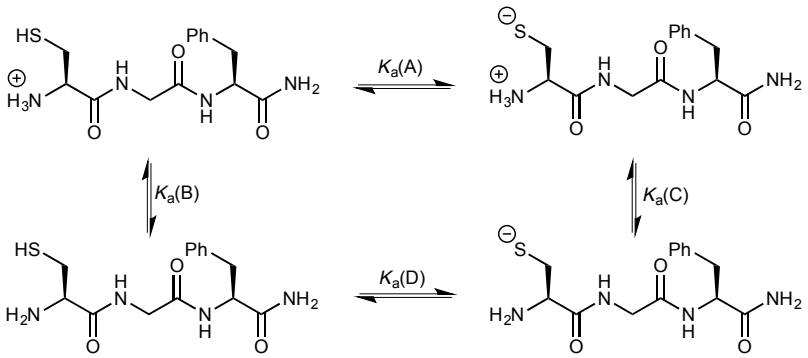
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50m M NaOH	12.40	1.056	1.004 [b]	1.00
40 mM NaOH	11.47	1.063	1.011	1.01
90% fb 0.1 M K_2CO_3	10.66	1.030	0.977	0.97
70% fb 0.1 M K_2CO_3	10.19	0.944	0.891	0.89
50% fb 0.1 M K_2CO_3	9.84	0.792	0.739	0.74
30% fb 0.1 M K_2CO_3	9.49	0.642	0.590	0.59
10% fb 0.1 M K_2CO_3	8.80	0.298	0.245	0.24
90% fb 0.1 M NH_4Cl	9.69	0.783	0.730	0.73
70% fb 0.1 M NH_4Cl	9.68	0.782	0.729	0.73
50% fb 0.1 M NH_4Cl	9.31	0.652	0.599	0.60
30% fb 0.1 M NH_4Cl	8.97	0.446	0.393	0.39
10% fb 0.1 M NH_4Cl	8.32	0.220	0.167	0.17
90% fb 0.1 M KH_2PO_4	7.66	0.164	0.111	0.11
70% fb 0.1 M KH_2PO_4	7.09	0.127	0.074	0.07
50% fb 0.1 M KH_2PO_4	6.72	0.082	0.029	0.03
30% fb 0.1 M KH_2PO_4	6.31	0.056	0.003	0.00
10% fb 0.1 M KH_2PO_4	5.69	0.047	-0.006	-0.01
30 mM HCl	1.75	0.049	-0.004	0.00
50 mM HCl	1.44	0.053 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S27: The variation of the fraction of thiol in thiolate form, f_{RS^-}, of pH for H-Cys-Gly-OH (**5**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S2.6 Determination of dissociation constants K_a(A) – K_a(D) for H-Cys-Gly-Phe-NH_2 (6)



The peptide H-Cys-Gly-Phe-NH_2 (**6**) was successfully prepared using solid phase peptide synthesis (See Experimental S2.6.1 for details).

A 0.1 M solution of H-Cys-Gly-Phe-NH_2 was prepared from 3.24 mg (9.99 x 10^{-6} M) of H-Cys-Gly-Phe-NH_2 and 100 µL of deionised water. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for H-Cys-Gly-Phe-NH_2 (Fig. S28 + S29, Table S10).

Values of dissociation constants, K_a(A) = 7.46 × 10^{-8} M, K_a(B) = 3.15 × 10^{-8} M, K_a(C) = 3.80 × 10^{-9} M, K_a(D) = 9.05 × 10^{-10} M were determined under reducing conditions for H-Cys-Gly-Phe-NH_2 by fitting of the data in Fig. S30 to Eq.2 (main text). K_a(C) was determined using Eq. S1.

Fig. S28: UV-Vis spectra for 0.2 mM H-Cys-Gly-Phe-NH_2 (**6**) in buffer solutions covering the pH range 1.58 – 12.40 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

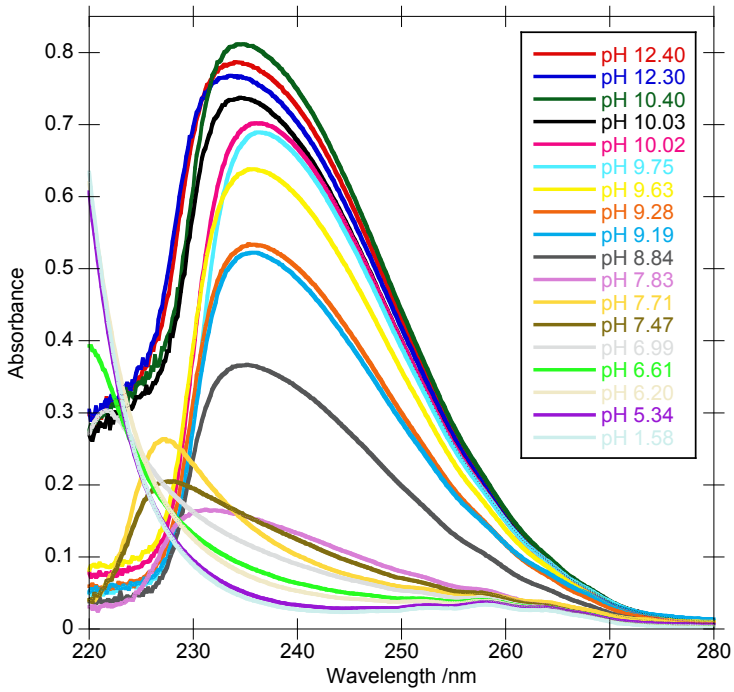


Fig. S29: UV-Vis spectra monitored at a single wavelength (λ = 230 – 236 nm) for 0.2 mM H-Cys-Gly-Phe-NH_2 (**6**) in buffer solutions covering the pH range 1.36 – 12.40 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

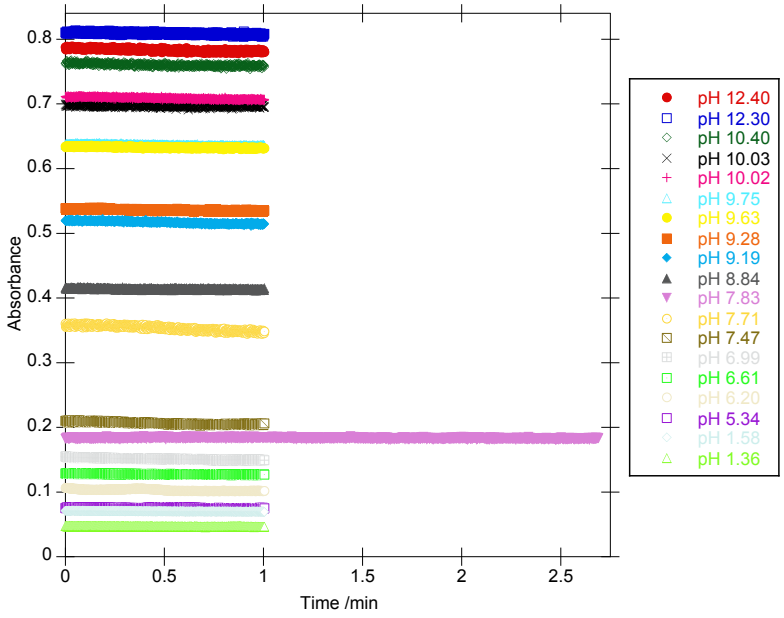


Fig. S30: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for H-Cys-Gly-Phe-NH_2 (**6**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).

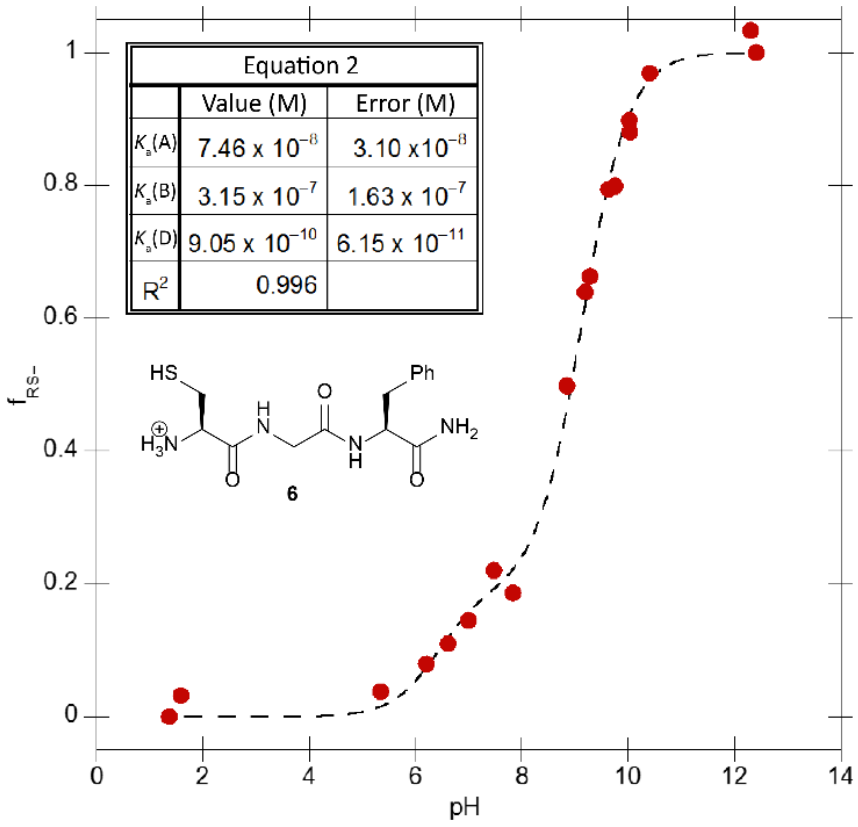
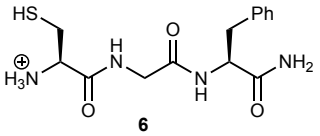


Table S10: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for H-Cys-Gly-Phe-NH_2 (**6**) at 25 °C and ionic strength I = 0.3 M (NaCl).

Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50m M NaOH	12.40	0.787	0.739 [b]	1.00
40 mM NaOH	12.30	0.811	0.764	1.03
90% fb 0.1 M K_2CO_3	10.40	0.764	0.716	0.97
70% fb 0.1 M K_2CO_3	10.03	0.698	0.651	0.88
90% fb 0.1 M NH_4Cl	10.02	0.711	0.664	0.90
50% fb 0.1 M K_2CO_3	9.75	0.638	0.591	0.80
70% fb 0.1 M NH_4Cl	9.63	0.634	0.587	0.79
50% fb 0.1 M NH_4Cl	9.28	0.538	0.491	0.66
30% fb 0.1 M K_2CO_3	9.19	0.520	0.472	0.64
30% fb 0.1 M NH_4Cl	8.84	0.416	0.368	0.50
10% fb 0.1 M NH_4Cl	7.83	0.185	0.137	0.19
10% fb 0.1 M K_2CO_3	7.71	0.359	0.312	0.42
90% fb 0.1 M KH_2PO_4	7.47	0.210	0.162	0.22
70% fb 0.1 M KH_2PO_4	6.99	0.155	0.108	0.15
50% fb 0.1 M KH_2PO_4	6.61	0.129	0.081	0.11
30% fb 0.1 M KH_2PO_4	6.20	0.107	0.059	0.08
10% fb 0.1 M KH_2PO_4	5.34	0.076	0.028	0.04
30 mM HCl	1.58	0.071	0.024	0.03
50 mM HCl	1.36	0.047 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

S2.6.1 Synthesis of H-Cys-Gly-Phe-NH_2 (6)



6

Solid Phase Peptide Synthesis (SPPS) was used to prepare H-Cys-Gly-Phe-NH_2. Rink amide (0.82 mmol/g) resin was weighed (0.28 g, 0.23 mmol) into a SPPS plastic vessel containing a frit. DMF (3 mL) was added and the resin was left to swell for 1 h. The DMF was drained and 20% (v/v) piperidine/DMF solution (3 mL) was added and the SPPS tube was placed on a shaker for 5 min in order to activate the resin. The solution was drained and a further amount of 20% (v/v) piperidine/DMF solution (3 mL) was added and the SPPS tube placed on the shaker for 10 min and then drained and washed with DMF.

Fmoc-Phe-OH (0.23 g, 0.60 mmol), PyBOP (0.31 g, 0.60 mmol) and of *N*,*N*-diisopropylethylamine (DIPEA) (0.21 mL, 1.21 mmol) dissolved in DMF (3 mL) were added to the resin and the flask placed on the shaker for 1 h. The solution was then drained and the resin washed with DMF (3 × 3 mL). This process was repeated to ensure all sites on the rink amide resin had Fmoc-Phe attached.

The Fmoc group was deprotected by the addition of 20 % (v/v) piperidine/DMF solution (3 mL) and placed on the shaker for 5 min. The solution was drained and a further amount of 20 % (v/v) piperidine/DMF solution (3 mL) was added. The SPPS tube was placed on the shaker for 10 min and then drained and washed with DMF (3 mL).

(0.18 g, 0.61 mmol) of Fmoc-Gly-OH, 0.34 g (0.66 mmol) of PyBOP and 0.21 mL (1.21 mmol) of DIPEA dissolved in DMF (3 mL) was added to the resin and then placed on the shaker for 1 h. The solution was then drained and the resin washed 3 times with DMF (3 mL). This process was repeated to ensure all peptide chains had Fmoc-Gly attached.

The Fmoc group was deprotected by the addition of 20 % (v/v) piperidine/DMF solution (3 mL) and placed on the shaker for 5 min. The solution was drained and a further amount of 20 % (v/v) piperidine/DMF solution (3 mL) was added and the SPPS tubeplaced on the shaker for 10 min and then drained and washed with DMF (3 mL).

The resin was then divided in two and half carried forward to prepare H-Cys-Gly-Phe-NH_2 and the other half to prepare H-Pen-Gly-Phe-NH_2 (see Section S2.7.1).

Fmoc-Cys(Trt)-OH (0.06 g, 0.11 mmol), PyBOP (0.18 g, 0.34 mmol) and DIPEA (0.10 mL, 0.57 mmol) were dissolved in DMF (3 mL) and added to the resin and placed on the shaker for 1 h. The solution was then drained and the resin washed with DMF (3 × 3 mL). This process was repeated to ensure all peptide chains had Fmoc-Cys(Trt) attached.

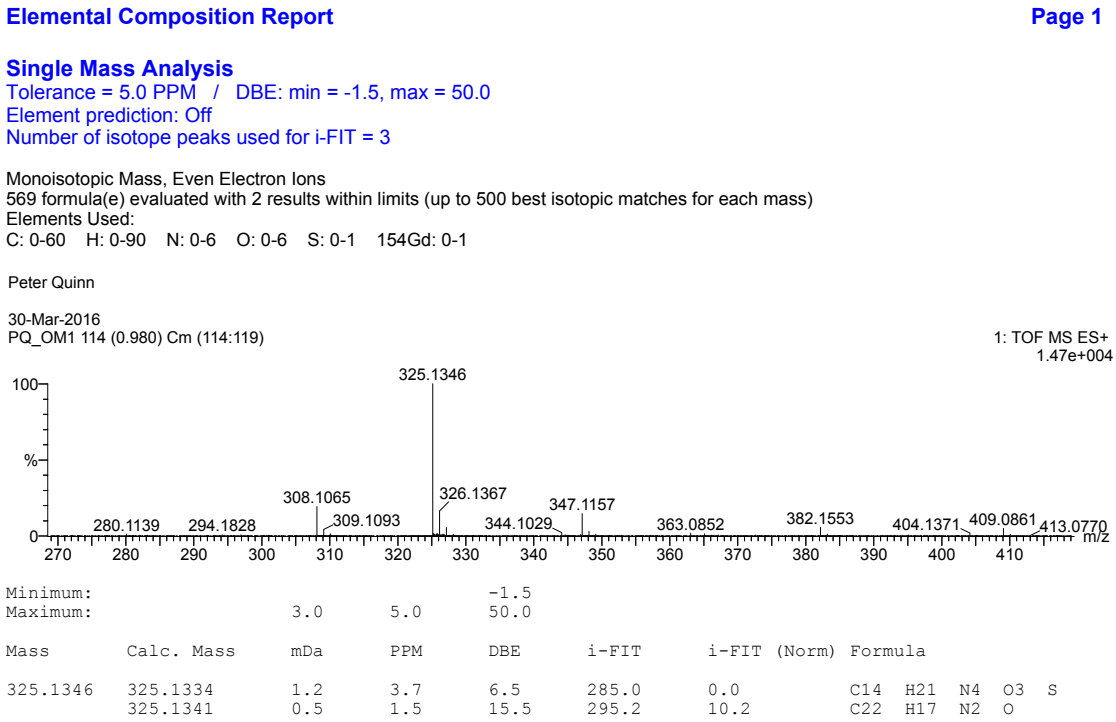
The Fmoc-deprotection was achieved by the addition of 20 % (v/v) piperidine/DMF solution (3 mL) and the SPPS tube was placed on shaker for 5 min. The solution was drained and a further amount of 20 % (v/v) piperidine/DMF solution (3 mL) was added and the tube placed on the shaker for 10 min and then drained and washed with DMF (3 mL).

The resin was then shrunk by the addition of diethyl ether (3 mL) and shaking for 1 h. The diethyl ether was then drained and the resin washed with diethyl ether (3 × 3 mL). The peptide was cleaved from the resin via the addition of 95 % (v/v) trifluoroacetic acid, 2.5 % water and 2.5 % triisopropylsilane (3 mL) then placed on the shaker for 4 h. The solvent was removed under vacuum at 50 °C to leave the peptide as a yellow solid.

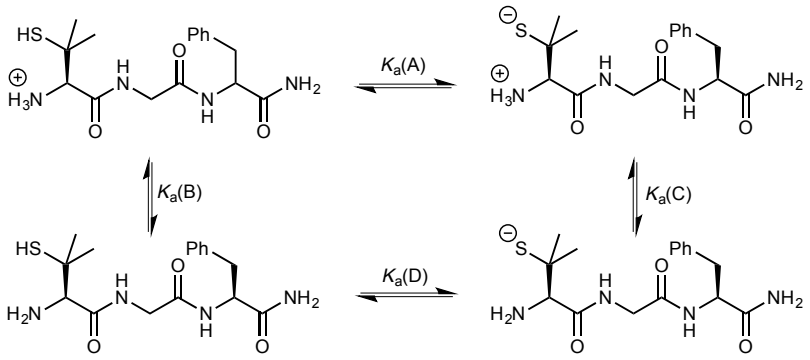
The peptide was purified using preparative HPLC. The peptide was dissolved in a 9 : 1 H_2O : MeOH solution (1.5 mL). The peptide was isolated using a gradient of 95 : 5 H_2O : MeOH to 50 : 50 H_2O : MeOH with a flow rate of 3 mL min^{-1} on a Sunfire C18 column attached to a Waters HPLC instrument. The peptide eluted with a retention time of 7 – 10 min.

The solvent was removed via freeze-drying to give the isolated H-Cys-Gly-Phe-NH_2 peptide as a white solid (15 mg, 4.6 × 10^{-5} mol, 41 % yield) ***m/z*** (ES+): 276 ($[M–H]^+$, 100%); **HRMS** (ES+): $[M–H]^+$ $C_{14}H_{21}N_4O_3S$ requires 325.1334, found 325.1346 (Fig. S31).

Fig. S31: High resolution mass spectrum of H-Cys-Gly-Phe-NH_2 (**6**).



S2.7 Determination of dissociation constants K_a(A) – K_a(D) H-Pen-Gly-Phe-NH_2 (7)



The peptide H-Pen-Gly-Phe-NH_2 (**7**) was successfully prepared using solid phase peptide synthesis (See Experimental S2.7.1 for details).

A 0.1 M solution of H-Pen-Gly-Phe-NH_2 was prepared from 3.52 mg (1.00 x 10^{-5} mol) of H-Pen-Gly-Phe-NH_2 and 100 µL of deionised water. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for H-Pen-Gly-Phe-NH_2 (Fig. S31 + S32, Table S11).

Values of dissociation constants, K_a(A) = 3.60 × 10^{-8} M, K_a(B) = 4.45 × 10^{-7} M, K_a(C) = 4.07 × 10^{-9} M, K_a(D) = 3.27 × 10^{-10} M were determined under reducing conditions for H-Pen-Gly-Phe-NH_2 by fitting of the data in Fig. S34 to Eq.2 as described above. K_a(C) was determined using Eq. S1.

Fig. S32: UV-Vis spectra for 0.2 mM H-Pen-Gly-Phe-NH_2 (**7**) in buffer solutions covering the pH range 1.54 – 12.39 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

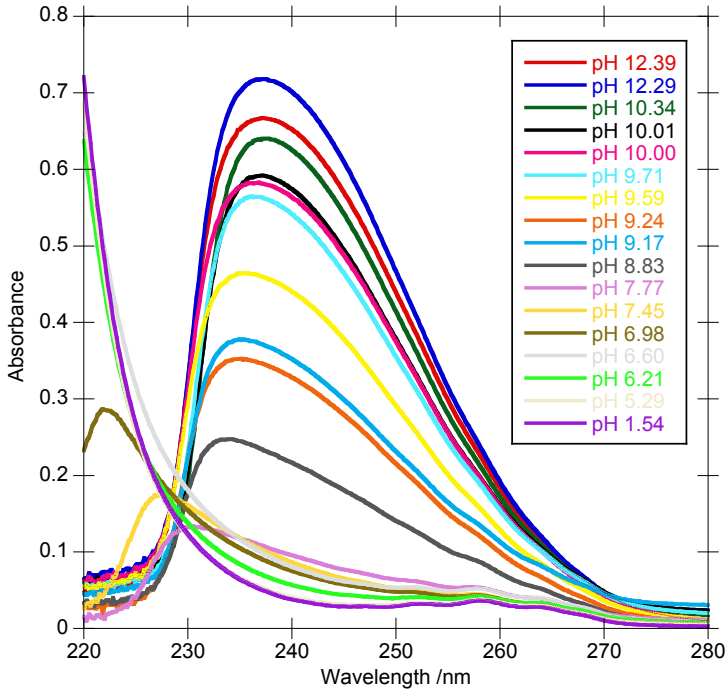


Fig. S33: UV-Vis spectra monitored at a single wavelength (λ = 230 – 236 nm) for 0.2 mM H-Pen-Gly-Phe-NH_2 (**7**) in buffer solutions covering the pH range 1.31 – 12.39 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

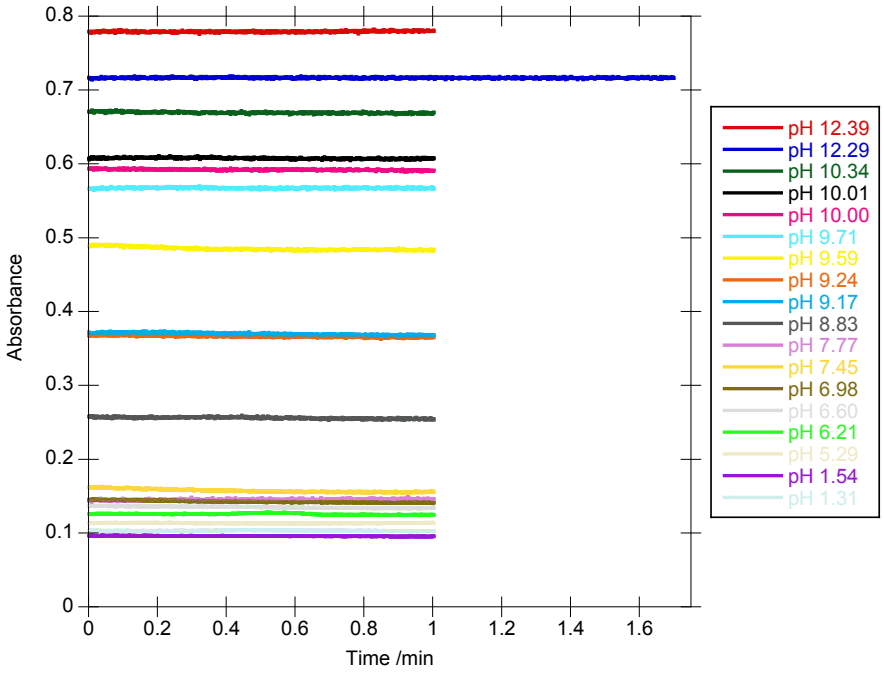
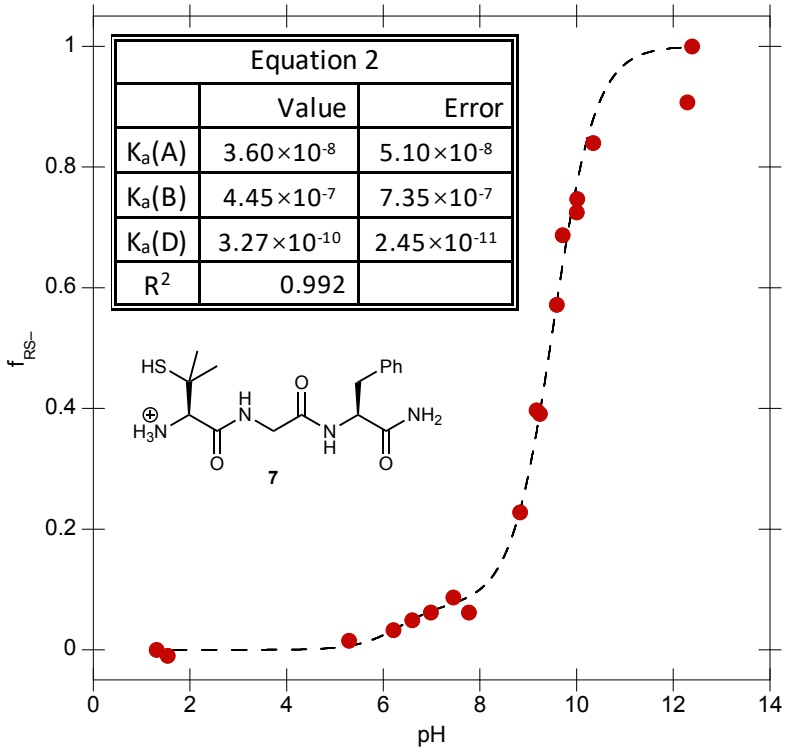


Table S11: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for H-Pen-Gly-Phe-NH_2 (**7**) at 25 °C and ionic strength I = 0.3 M (NaCl).

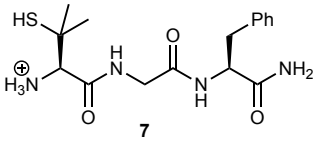
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50m M NaOH	12.39	0.779	0.676 [b]	1.00
40 mM NaOH	12.29	0.717	0.613	0.91
90% fb 0.1 M K_2CO_3	10.34	0.671	0.568	0.84
70% fb 0.1 M K_2CO_3	10.01	0.609	0.505	0.75
50% fb 0.1 M K_2CO_3	9.71	0.568	0.464	0.69
30% fb 0.1 M K_2CO_3	9.17	0.372	0.268	0.40
90% fb 0.1 M NH_4Cl	10.00	0.594	0.490	0.73
70% fb 0.1 M NH_4Cl	9.59	0.490	0.386	0.57
50% fb 0.1 M NH_4Cl	9.24	0.368	0.264	0.39
30% fb 0.1 M NH_4Cl	8.83	0.258	0.154	0.23
10% fb 0.1 M NH_4Cl	7.77	0.145	0.042	0.06
90% fb 0.1M KH_2PO_4	7.45	0.162	0.059	0.09
70% fb 0.1 M KH_2PO_4	6.98	0.145	0.042	0.06
50% fb 0.1 M KH_2PO_4	6.60	0.137	0.033	0.05
30% fb 0.1 M KH_2PO_4	6.21	0.126	0.023	0.03
10% fb 0.1 M KH_2PO_4	5.29	0.114	0.010	0.02
30 mM HCl	1.54	0.097	-0.007	-0.01
50 mM HCl	1.31	0.103 [a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S34: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for H-Pen-Gly-Phe-NH_2 (**7**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S2.7.1 Synthesis of H-Pen-Gly-Phe-NH_2 (7)



The H-Gly-Phe-NH-Resin was prepared as described for H-Cys-Gly-Phe-NH_2 **6**. Fmoc-Pen(Trt)-OH (0.06 g, 0.01 mmol), PyBOP (0.17 g, 0.33 mmol) and DIPEA (0.10 mL, 0.57 mmol) dissolved in DMF (3 mL) was added to the resin in the SPPS tube and then placed on the shaker for 1 h. The solution was then drained and the resin washed with DMF (3 × 3 mL). This process was repeated to ensure all peptide chains had Fmoc-Pen(Trt) attached.

The Fmoc was deprotected by the addition of 20 % (v/v) piperidine/DMF solution (3 mL) and placed on the shaker for 5 min. The solution was drained and a further amount of 20 % (v/v) piperidine/DMF solution (3 mL) was added and the SPPS tube placed on the shaker for 10 min and then drained and washed with DMF (3 mL).

The resin was then shrunk by the addition of diethyl ether (3 mL) and shaking for 1 h. The diethyl ether was drained and the resin washed 3 × with diethyl ether (3 mL). The peptide was

then cleaved from the resin via the addition of 95 % (v/v) trifluoroacetic acid, 2.5 % water and 2.5 % triisopropylsilane (3 mL) and placing on the shaker for 4 h. The solvent was then removed under vacuum at 50 °C to leave the peptide as a yellow solid.

The peptide was purified using preparative HPLC. The peptide was dissolved in of 9 : 1 H_2O : MeOH solution (1.5 mL). The peptide was isolated using a gradient of 95 : 5 H_2O : MeOH to 50 : 50 H_2O : MeOH with a flow rate of 3 mL min^{-1} on a Sunfire C18 column attached to a Waters HPLC machine. The peptide eluted with a retention time of 7 – 10 min.

The solvent was removed via freeze-drying to give the isolated H-Pen-Gly-Phe-NH_2 **7** peptide as a white solid (12 mg, 3.4 × 10^{-5} mol, 30 % yield) ***m/z*** (ES+): 353 ($[M–H]^+$, 100%); **HRMS** (ES+): $[M–H]^+$ $C_{16}H_{25}N_4O_3S$ requires 353.1647, found 353.1644 (Fig. S35).

Fig. S35: High resolution mass spectrum of H-Pen-Gly-Phe-NH_2 (**7**).

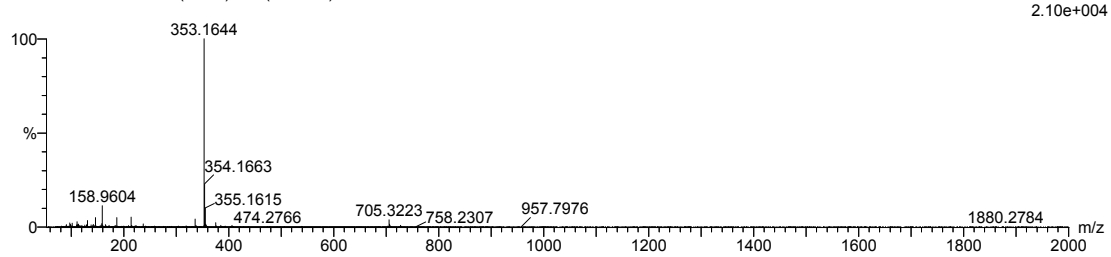
Elemental Composition Report **Page 1**

Single Mass Analysis
Tolerance = 3.0 mDa / DBE: min = -1.5, max = 50.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 3

Monoisotopic Mass, Even Electron Ions
4418 formula(e) evaluated with 17 results within limits (up to 50 best isotopic matches for each mass)
Elements Used:
C: 0-50 H: 0-100 N: 0-10 O: 0-10 P: 0-1 S: 0-3 96Ru: 0-1

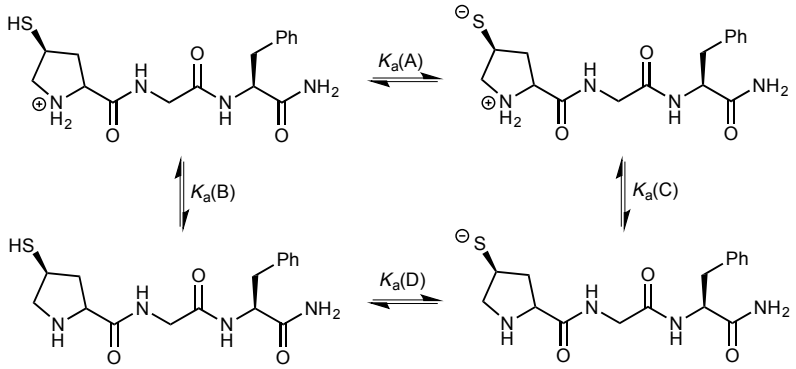
Oliver Maguire
OMxPEPx002x007 175 (1.446) Cm (175:179)

1: TOF MS ES+
2.10e+004



Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	i-FIT (Norm)	Formula
Minimum:				-1.5			
Maximum:		3.0	5.0	50.0			
353.1644	353.1654	-1.0	-2.8	2.5	327.4	0.6	C9 H25 N10 O S2
	353.1637	0.7	2.0	2.5	328.7	1.9	C10 H26 N8 O2 P S
	353.1647	-0.3	-0.8	6.5	329.0	2.2	C16 H25 N4 O3 S
	353.1643	0.1	0.3	0.5	329.5	2.7	C16 H33 O2 S3
	353.1621	2.3	6.5	7.5	329.6	2.7	C12 H21 N10 O S
	353.1634	1.0	2.8	1.5	330.4	3.6	C15 H29 O7 S
	353.1616	2.8	7.9	1.5	330.9	4.1	C12 H29 N6 S3
	353.1664	-2.0	-5.7	1.5	331.1	4.3	C14 H30 N2 O4 P S
	353.1630	1.4	4.0	6.5	337.9	11.1	C17 H26 N2 O4 P
	353.1644	0.0	0.0	11.5	337.9	11.1	C18 H22 N6 P
	353.1654	-1.0	-2.8	15.5	338.0	11.1	C24 H21 N2 O
	353.1670	-2.6	-7.4	10.5	338.1	11.3	C22 H26 O2 P
	353.1672	-2.8	-7.9	2.5	338.3	11.5	C12 H25 N4 O8
	353.1646	-0.2	-0.6	3.5	338.4	11.6	C8 H21 N10 O6
	353.1632	1.2	3.4	-1.5	338.7	11.9	C7 H25 N6 O10
	353.1669	-2.5	-7.1	0.5	338.9	12.1	C15 H33 N2 O 96Ru
	353.1662	-1.8	-5.1	-1.5	339.1	12.3	C6 H26 N8 O7 P

S2.8 Determination of dissociation constants K_a(A) – K_a(D) H-(4*S*)-Mcp-Gly-Phe-NH_2 (8)



The peptide H-(4*S*)-Mcp-Gly-Phe-NH_2 (**8**) was successfully prepared using solid phase peptide synthesis (See Experimental S2.8.1 for details).

A 0.01 M solution of H-(4*S*)-Mcp-Gly-Phe-NH_2 was prepared from 35.0 mg (1.00 x 10^{-4} mol) of H-(4*S*)-Mcp-Gly-Phe-NH_2 in a 10 mL volumetric flask. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for H-(4*S*)-Mcp-Gly-Phe-NH_2 (Fig. S35 + S36, Table S12).

Values of dissociation constants, K_a(A) = 4.36 × 10^{-8} M, K_a(B) = 6.59 × 10^{-8} M, K_a(C) = 1.95 × 10^{-9} M, K_a(D) = 1.29 × 10^{-9} M were determined under reducing conditions for H-Pen-Gly-Phe-NH_2 by fitting of the data in Fig. S38 to Eq.2 as described above. K_a(C) was determined using Eq. S1.

Fig. S36: UV-Vis spectra for 0.2 mM H-(4*S*)-Mcp-Gly-Phe-NH_2 (**8**) in buffer solutions covering the pH range 1.38 – 12.55 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

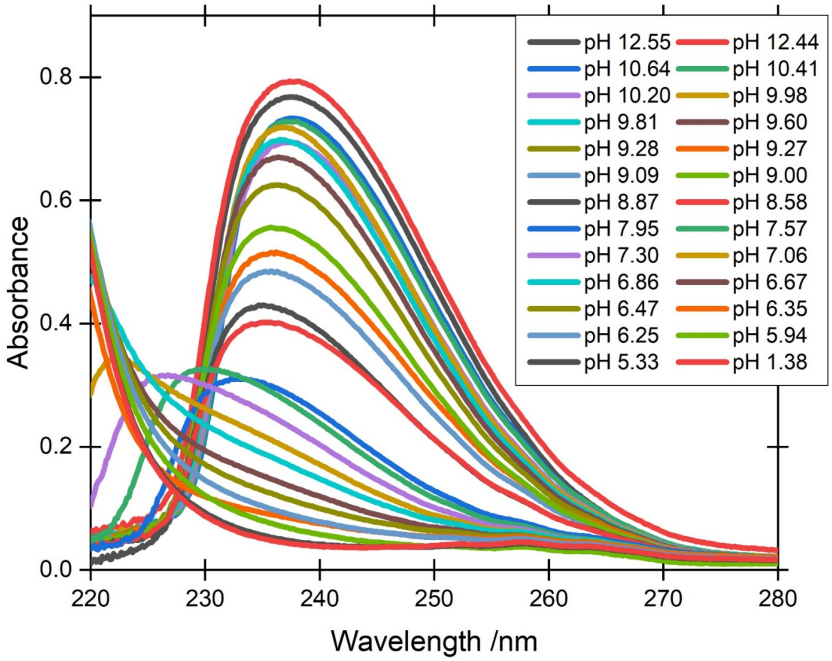
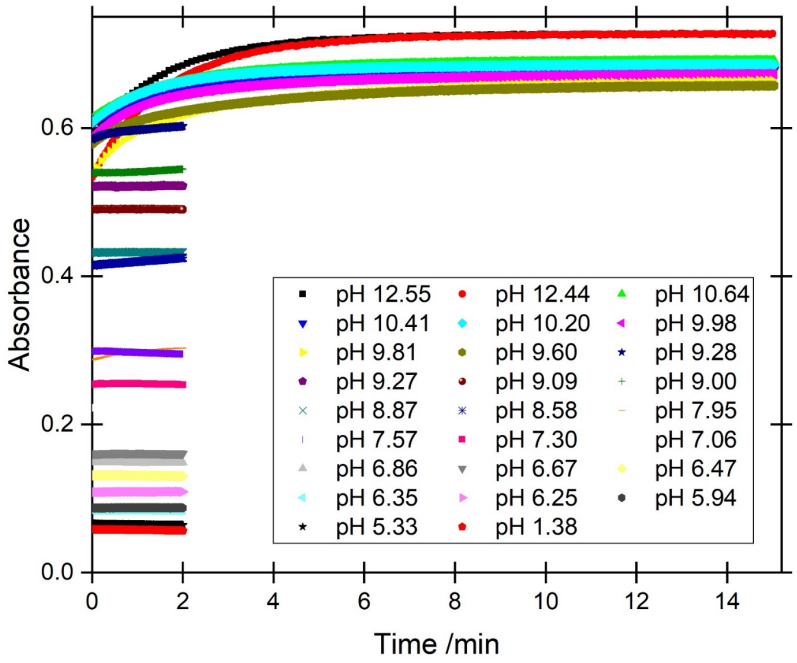


Fig. S37: UV-Vis spectra monitored at a single wavelength (λ = 230 – 238 nm) for 0.2 mM H-(4*S*)-Mcp-Gly-Phe-NH_2 (**8**) in buffer solutions covering the pH range 1.38 – 12.55 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).



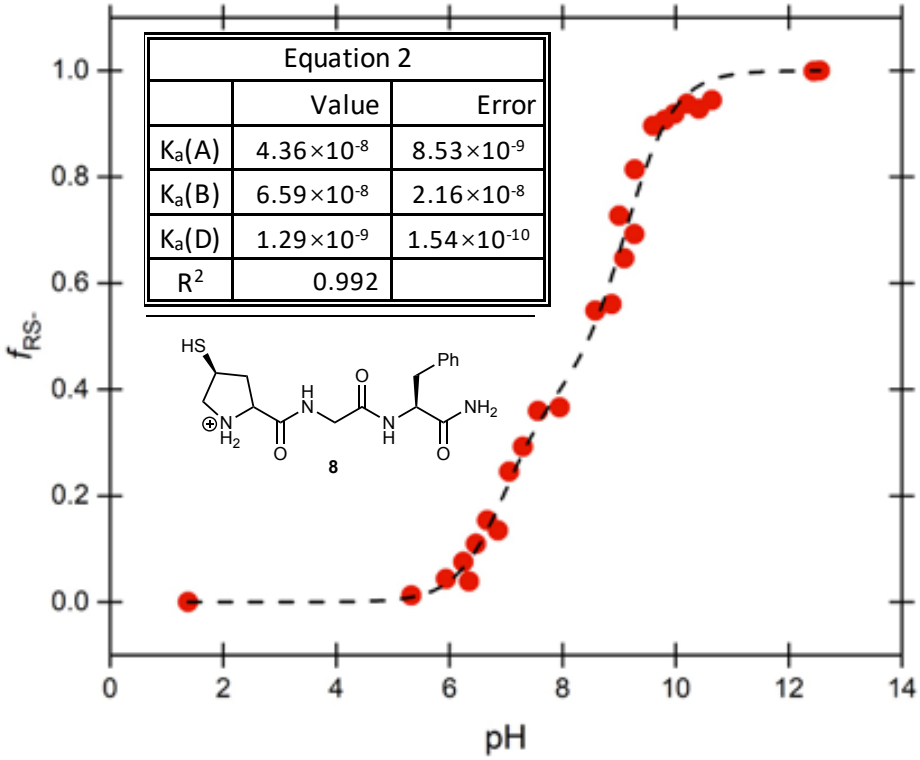
The rise in absorbance over time observed at higher pHs in Fig. S37 indicates that either the H-(4*S*)-Mcp-Gly-Phe-NH_2 (**8**) had oxidised in the solid form or that the H-(4*S*)-Mcp-Gly-Phe-NH_2 had oxidised in solution prior to the addition of TCEP solution. To ensure that full reduction of disulfide bonds occurred a fresh TCEP solution was prepared for every measurement from solid TCEP.HCl and the appropriate buffer. The rise in absorbance as the experiment progressed is a result of the TCEP reducing the disulfide bonds. After rising the absorbance then plateaus indicative of full reduction of the disulfide bonds and a consistent thiol concentration.

Table S12: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for H-(4*S*)-Mcp-Gly-Phe-NH_2 (**8**) at 25 °C and ionic strength I = 0.3 M (NaCl).

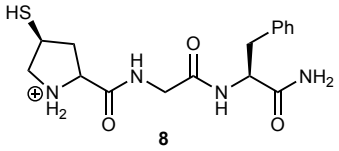
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.55	0.727	0.669[b]	1.00
40 mM NaOH	12.44	0.726	0.668	1.00
90% fb 0.1 M K_2CO_3	10.64	0.689	0.631	0.94
80% fb 0.1 M K_2CO_3	10.41	0.679	0.621	0.93
70% fb 0.1 M K_2CO_3	10.20	0.685	0.627	0.94
60% fb 0.1 M K_2CO_3	9.98	0.672	0.614	0.92
50% fb 0.1 M K_2CO_3	9.81	0.664	0.606	0.91
40% fb 0.1 M K_2CO_3	9.60	0.658	0.600	0.90
30% fb 0.1 M K_2CO_3	9.28	0.603	0.544	0.81
50% fb 0.1 M NH_4Cl	9.27	0.521	0.463	0.69
40% fb 0.1 M NH_4Cl	9.09	0.490	0.432	0.65
20% fb 0.1 M K_2CO_3	9.00	0.544	0.486	0.73
30% fb 0.1 M NH_4Cl	8.87	0.433	0.375	0.56
10% fb 0.1 M K_2CO_3	8.58	0.425	0.367	0.55
20% fb 0.1 M NH_4Cl	7.95	0.303	0.245	0.37
90% fb 0.1M KH_2PO_4	7.57	0.299	0.241	0.36
80% fb 0.1M KH_2PO_4	7.30	0.254	0.196	0.29
70% fb 0.1M KH_2PO_4	7.06	0.222	0.164	0.25
60% fb 0.1M KH_2PO_4	6.86	0.148	0.090	0.13
50% fb 0.1M KH_2PO_4	6.67	0.161	0.103	0.15
40% fb 0.1 M KH_2PO_4	6.47	0.131	0.073	0.11
10% fb 0.1 M NH_4Cl	6.35	0.084	0.026	0.04
30% fb 0.1 M KH_2PO_4	6.25	0.108	0.050	0.08
20% fb 0.1 M KH_2PO_4	5.94	0.087	0.029	0.04
10% fb 0.1M KH_2PO_4	5.33	0.066	0.008	0.01
50 mM HCl	1.38	0.058[a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S38: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for H-(4*S*)-Mcp-Gly-Phe-NH_2 (**8**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S2.8.1 Synthesis of H-(4*S*)-Mcp-Gly-Phe-NH_2 (8)



As with the other tripeptides this compound was synthesised using Fmoc solid phase techniques. Rink amide MBHA resin (100-200 mesh, 0.436 mmol/g) (0.69 g, 0.30 mmol) was swollen for 1 h in DMF (5 mL). The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Phe-OH (0.29 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin. The resin was agitated for 1 h, drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Gly-OH (0.22 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin. The

resin was agitated for 1 h, drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-L-Pro(4-STrt)-OH (2*S*,4*S*) (0.37 g, 0.6 mmol), PyBOP (0.31 g, 0.60 mmol) and DIPEA (0.16 g, 1.20 mmol) were dissolved in DMF (3 mL) and added to the resin. The resin was agitated for 2 h drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). The resin was treated with Et_2O (5 mL, 30 min) drained and washed with Et_2O (2 × 5 mL). A solution of TFA, triisopropylsilane and water 95/2.5/2.5 (v/v) (3 mL) was added to the resin and agitated for 1 h. The solution was recovered and the resin washed with DCM (3 × 5 mL), the combined organic solutions were concentrated under reduced pressure. The peptide was subsequently purified by preparative HPLC using a gradient of 95:5:0.05 H_2O:MeCN:TFA to 5:95:0.03 H_2O:MeCN:TFA over 120 min with a flow rate of 2 mL min^{-1} on a Discovery BIO Wide Pore C18-5 column with a retention time of 28-34 min. The solvent was removed *via* freeze-drying to give the isolated peptide as a white solid (97 mg, 0.27 mmol, 90%). ***m/z*** (ES+): 351 ($[M+H]^+$, 100%); **HRMS** (ES+): $[M+H]^+$ $C_{16}H_{23}N_4O_3S$ requires 351.1491, found 351.1480 (Fig. S39).

Fig. S39: High resolution mass spectrum of H-(4*S*)-Mcp-Gly-Phe-NH_2 (**8**).

Elemental Composition Report **Page 1**

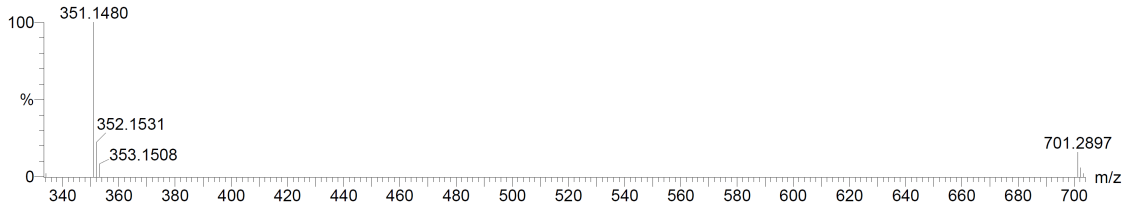
Single Mass Analysis
Tolerance = 3.0 mDa / DBE: min = -1.5, max = 100.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 5

Monoisotopic Mass, Even Electron Ions
1618 formula(e) evaluated with 16 results within limits (up to 500 closest results for each mass)
Elements Used:
C: 0-40 H: 0-80 N: 0-6 O: 0-6 P: 0-2 S: 0-2

QToF Premier

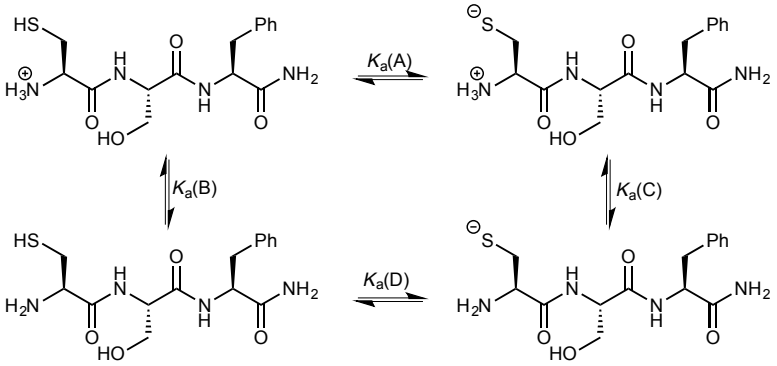
24-Oct-2019
WB6-4 202 (1.707)

1: TOF MS ES+
3.93e+004



Minimum:				-1.5			
Maximum:		3.0	5.0	100.0			
Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	i-FIT (Norm)	Formula
351.1480	351.1484	-0.4	-1.1	-1.5	45.3	9.2	C8 H27 N6 O5 S2
	351.1474	0.6	1.7	7.5	39.5	3.4	C17 H24 N2 O4 P
	351.1487	-0.7	-2.0	12.5	37.4	1.3	C18 H20 N6 P
	351.1490	-1.0	-2.8	2.5	39.0	2.9	C15 H29 O5 P2
	351.1491	-1.1	-3.1	7.5	43.2	7.1	C16 H23 N4 O3 S
	351.1467	1.3	3.7	-1.5	43.3	7.2	C9 H28 N4 O6 P S
	351.1465	1.5	4.3	6.5	43.0	6.9	C19 H29 P2 S
	351.1497	-1.7	-4.8	-1.5	43.3	7.2	C8 H29 N6 O3 P2 S
	351.1497	-1.7	-4.8	16.5	40.9	4.9	C24 H19 N2 O
	351.1463	1.7	4.8	3.5	37.3	1.2	C11 H25 N6 O3 P2
	351.1499	-1.9	-5.4	1.5	45.2	9.1	C16 H33 P2 S2
	351.1457	2.3	6.5	12.5	40.1	4.0	C19 H19 N4 O3
	351.1504	-2.4	-6.8	7.5	37.3	1.2	C16 H25 N4 O P2
	351.1507	-2.7	-7.7	2.5	43.1	7.0	C14 H28 N2 O4 P S
	351.1452	2.8	8.0	6.5	45.4	9.3	C19 H27 O2 S2
	351.1451	2.9	8.3	3.5	44.0	7.9	C11 H23 N6 O5 S

S2.9 Determination of dissociation constants K_a(A) – K_a(D) H-Cys-Ser-Phe-NH_2 (9)



The peptide H-Cys-Ser-Phe-NH_2 (**9**) was successfully prepared using solid phase peptide synthesis (See Experimental S2.9.1 for details).

A 0.01 M solution of H-Cys-Ser-Phe-NH_2 was prepared from 36.6 mg (1.00×10^{-4} mol) of H-Cys-Ser-Phe-NH_2 in a 10 mL volumetric flask. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for H-Cys-Ser-Phe-NH_2 (Fig. S39 + S40, Table S13).

Values of dissociation constants, K_a(A) = 4.93×10^{-8} M, K_a(B) = 2.13×10^{-7} M, K_a(C) = 3.75×10^{-9} M, K_a(D) = 8.69×10^{-10} M were determined under reducing conditions for H-Cys-Ser-Phe-NH_2 by fitting of the data in Fig. S42 to Eq.2 as described above. K_a(C) was determined using Eq. S1.

Fig. S40: UV-Vis spectra for 0.2 mM H-Cys-Ser-Phe-NH_2 (**9**) in buffer solutions covering the pH range 1.38 – 12.53 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

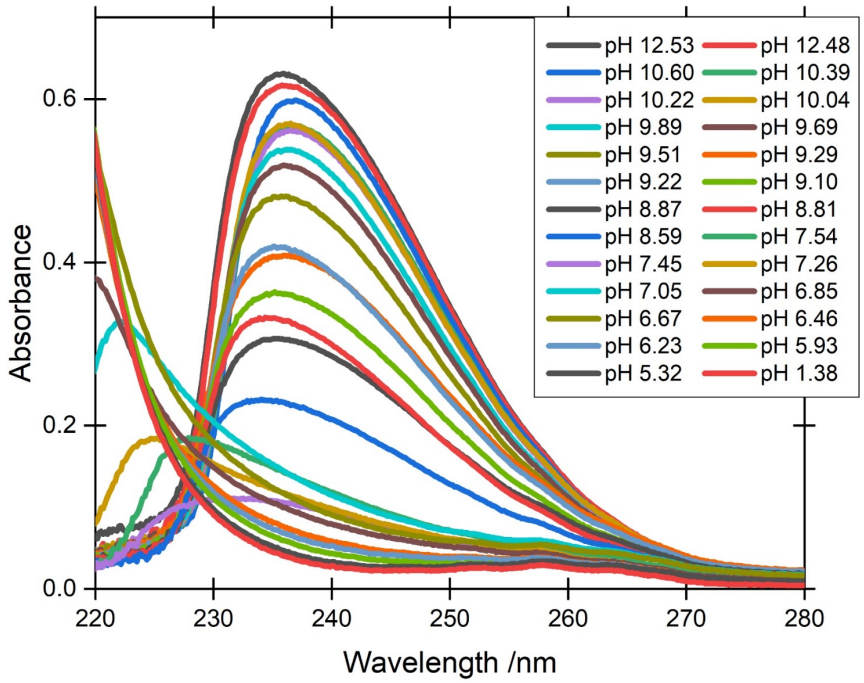


Fig. S41: UV-Vis spectra monitored at a single wavelength (λ = 229 – 237 nm) for 0.2 mM H-Cys-Ser-Phe-NH_2 (**9**) in buffer solutions covering the pH range 1.38 – 12.53 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

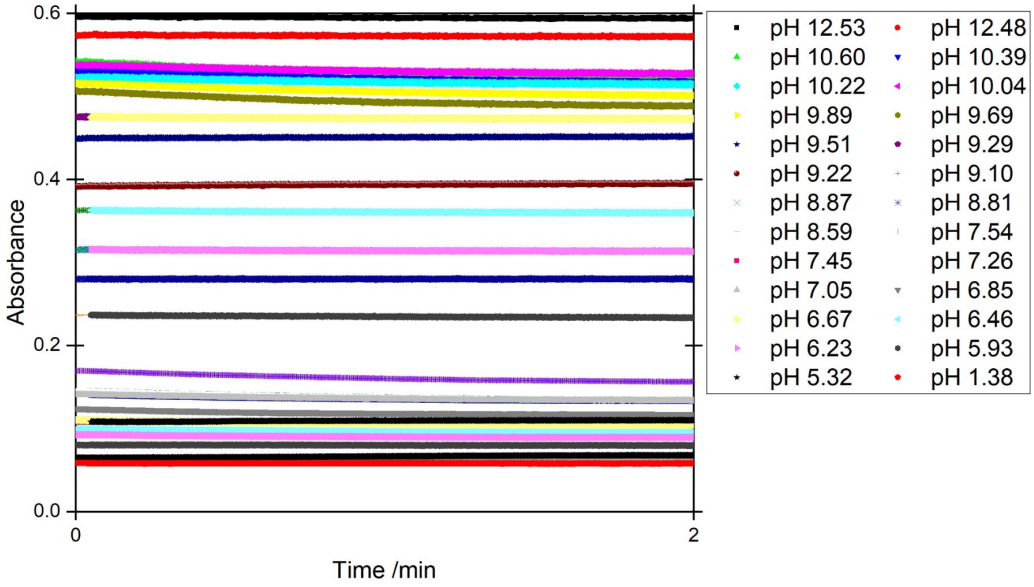
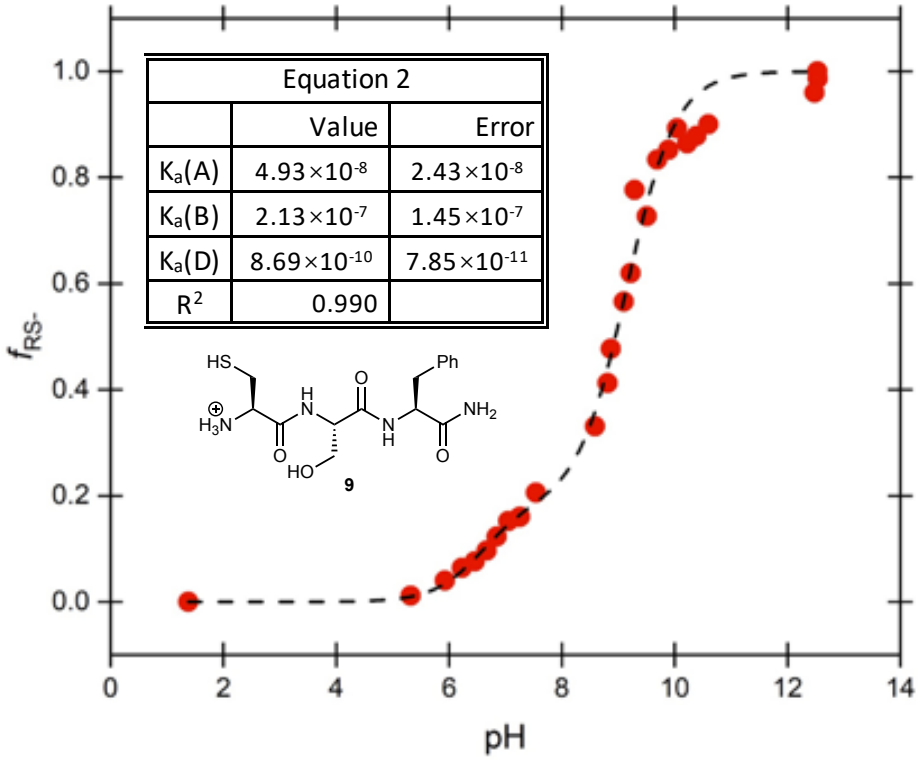


Table S13: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for H-Cys-Ser-Phe-NH_2 (**9**) at 25 °C and ionic strength I = 0.3 M (NaCl).

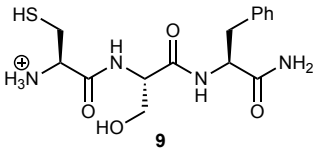
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.53	0.596	0.538[b]	1.00
40 mM NaOH	12.48	0.574	0.516	0.96
90% fb 0.1 M K_2CO_3	10.60	0.542	0.484	0.90
80% fb 0.1 M K_2CO_3	10.39	0.531	0.472	0.88
70% fb 0.1 M K_2CO_3	10.22	0.523	0.464	0.86
60% fb 0.1 M K_2CO_3	10.04	0.538	0.480	0.89
50% fb 0.1 M K_2CO_3	9.89	0.516	0.458	0.85
40% fb 0.1 M K_2CO_3	9.69	0.506	0.448	0.83
30% fb 0.1 M K_2CO_3	9.51	0.449	0.391	0.73
50% fb 0.1 M NH_4Cl	9.29	0.476	0.417	0.78
20% fb 0.1 M K_2CO_3	9.22	0.391	0.333	0.62
40% fb 0.1 M NH_4Cl	9.10	0.363	0.304	0.57
30% fb 0.1 M NH_4Cl	8.87	0.315	0.256	0.48
10% fb 0.1 M K_2CO_3	8.81	0.280	0.222	0.41
20% fb 0.1 M NH_4Cl	8.59	0.236	0.178	0.33
90% fb 0.1M KH_2PO_4	7.54	0.169	0.111	0.21
80% fb 0.1M KH_2PO_4	7.26	0.145	0.086	0.16
70% fb 0.1M KH_2PO_4	7.05	0.140	0.082	0.15
60% fb 0.1M KH_2PO_4	6.85	0.124	0.066	0.12
50% fb 0.1M KH_2PO_4	6.67	0.110	0.052	0.10
40% fb 0.1 M KH_2PO_4	6.46	0.099	0.041	0.08
30% fb 0.1 M KH_2PO_4	6.23	0.093	0.034	0.06
20% fb 0.1 M KH_2PO_4	5.93	0.080	0.022	0.04
10% fb 0.1M KH_2PO_4	5.32	0.065	0.006	0.01
50 mM HCl	1.38	0.058[a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S42: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for H-Cys-Ser-Phe-NH_2 (**9**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S2.9.1 Synthesis of H-Cys-Ser-Phe-NH_2 (9)



As with the other tripeptides this compound was synthesised using Fmoc solid phase techniques. Rink amide MBHA resin (100-200 mesh, 0.436 mmol/g) (0.69 g, 0.30 mmol) was swollen for 1 h in DMF (5 mL). The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Phe-OH (0.29 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin. The resin was agitated for 1 h, drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Ser(tBu)-OH (0.29 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin.

The resin was agitated for 1 h, drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Cys(mmt)-OH (0.46 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin. The resin was agitated for 2 h drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). The resin was treated with Et_2O (5 mL, 30 min) drained and washed with Et_2O (2 × 5 mL). A solution of TFA:triisopropylsilane:water 95:2.5:2.5 (v/v) (3 mL) was added to the resin and agitated for 1 h. The solution was recovered and the resin washed with DCM (3 × 5 mL), the combined organic solutions were concentrated under reduced pressure. The peptide was triturated with ether (2 × 10 mL). The peptide was subsequently purified by preparatory HPLC using a gradient of 95:5:0.05 H_2O:MeCN:TFA to 5:95:0.03 H_2O:MeCN:TFA over 120 min with a flow rate of 2.3 mL min^{-1} on a Discovery BIO Wide Pore C18-5 column with a retention time of 8-10 min. The solvent was removed via freeze-drying to give the isolated peptide as a white solid (78 mg, 0.22 mmol, 73% yield) ***m/z*** (ES+): 355 ($[M+H]^+$, 100%); **HRMS** (ES+): $[M+H]^+$ $C_{15}H_{23}N_4O_4S$ requires 355.1440, found 355.1442 (Fig. S43).

Fig. S43: High resolution mass spectrum of H-Cys-Ser-Phe-NH_2 (**9**).

Elemental Composition Report **Page 1**

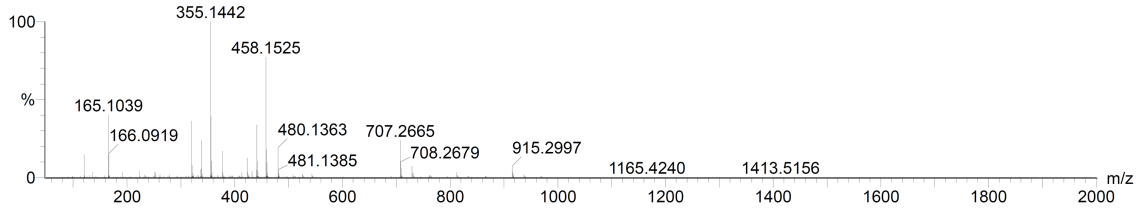
Single Mass Analysis
Tolerance = 3.0 mDa / DBE: min = -1.5, max = 100.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 5

Monoisotopic Mass, Even Electron Ions
898 formula(e) evaluated with 5 results within limits (up to 500 closest results for each mass)
Elements Used:
C: 0-60 H: 0-50 N: 0-8 O: 0-8 S: 0-2

QToF Premier

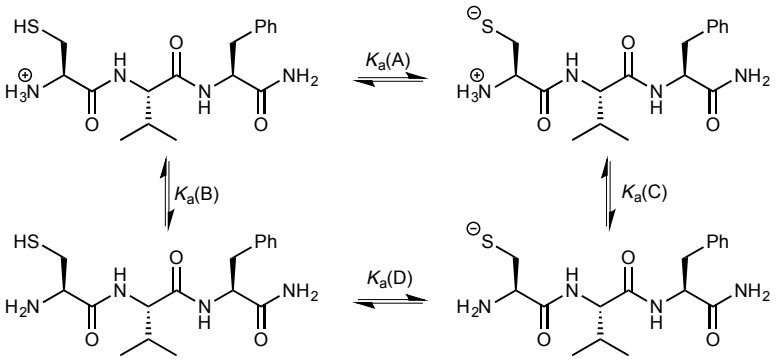
02-Oct-2019
WB5-146 198 (1.669) Cm (194:205)

1: TOF MS ES+
5.34e+004



Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	i-FIT (Norm)	Formula
Minimum:				-1.5			
Maximum:		3.0	5.0	100.0			
355.1442	355.1440	0.2	0.6	6.5	909.0	1.4	C15 H23 N4 O4 S
	355.1447	-0.5	-1.4	15.5	909.7	2.2	C23 H19 N2 O2
	355.1453	-1.1	-3.1	11.5	908.3	0.7	C16 H19 N8 S
	355.1427	1.5	4.2	1.5	909.7	2.2	C14 H27 O8 S
	355.1420	2.2	6.2	16.5	910.1	2.6	C19 H15 N8

S2.10 Determination of dissociation constants K_a(A) – K_a(D) H-Cys-Val-Phe-NH_2 (10)



The peptide H-Cys-Val-Phe-NH_2 (**10**) was successfully prepared using solid phase peptide synthesis (See Experimental S2.10.1 for details).

A 0.01 M solution of H-Cys-Val-Phe-NH_2 was prepared from 35.4 mg (1.00×10^{-4} mol) of H-Cys-Val-Phe-NH_2 in a 10 mL volumetric flask. An identical procedure to that for cysteine methyl ester (**2**) was used for determining the dissociation constants K_a(A) – K_a(D) for H-Cys-Val-Phe-NH_2 (Fig. S43 + S44, Table S14).

Values of dissociation constants, $K_a(A) = 4.12 \times 10^{-8}$ M, $K_a(B) = 1.80 \times 10^{-7}$ M, $K_a(C) = 4.06 \times 10^{-9}$ M, $K_a(D) = 9.29 \times 10^{-10}$ M were determined under reducing conditions for H-Cys-Val-Phe-NH_2 by fitting of the data in Fig. S46 to Eq.2 as described above. K_a(C) was determined using Eq. S1.

Fig. S44: UV-Vis spectra for 0.2 mM H-Cys-Val-Phe-NH_2 (**10**) in buffer solutions covering the pH range 1.41 – 12.54 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

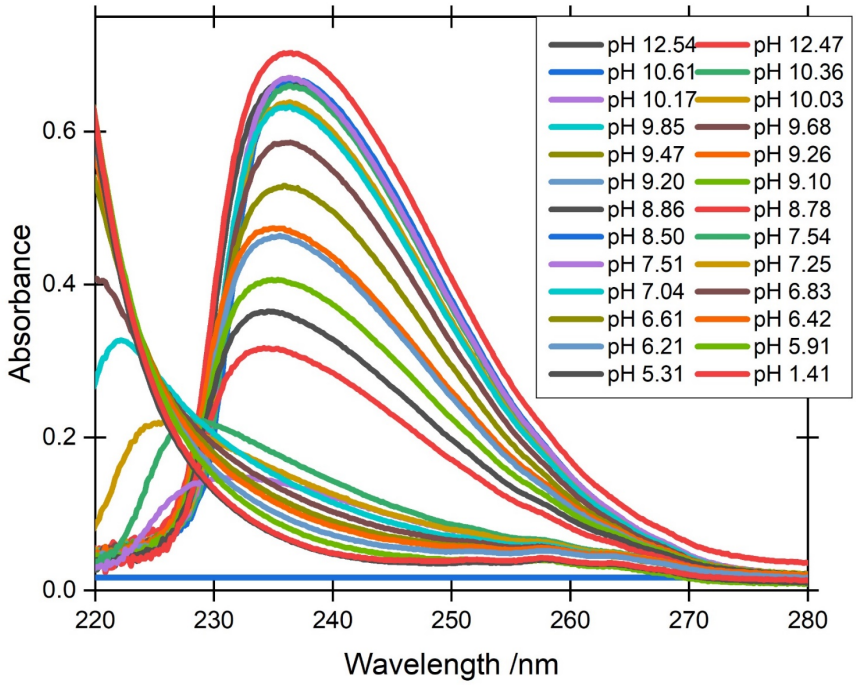


Fig. S45: UV-Vis spectra monitored at a single wavelength (λ = 229 – 237 nm) for 0.2 mM H-Cys-Val-Phe-NH_2 (**10**) in buffer solutions covering the pH range 1.41 – 12.54 with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

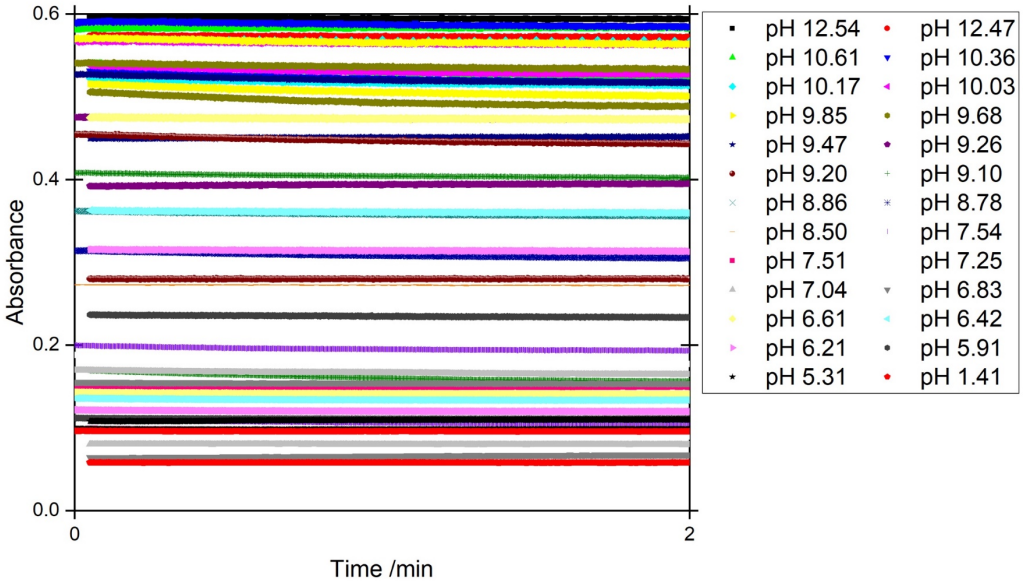
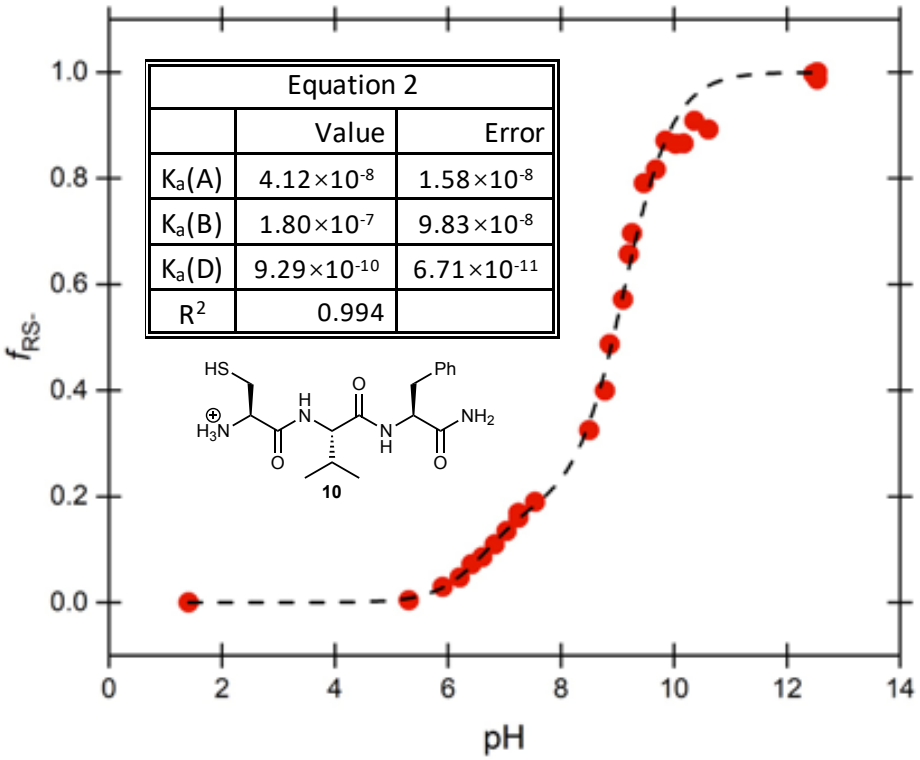


Table S14: A_{obs}, $A_{obs} - A_{237}^{RSH}$ and fraction of thiol in thiolate form, f_{RS^-}, at each pH for H-Cys-Val-Phe-NH_2 (**10**) at 25 °C and ionic strength I = 0.3 M (NaCl).

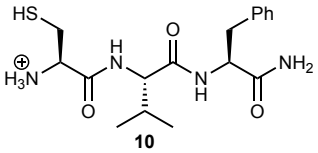
Solution	pH	A_{obs}	$A_{obs} - A_{237}^{RSH}$	f_{RS^-}
50 mM NaOH	12.54	0.641	0.545[b]	1.00
50 mM NaOH	12.54	0.634	0.538	0.99
40 mM NaOH	12.47	0.639	0.543	1.00
90% fb 0.1 M K_2CO_3	10.61	0.582	0.486	0.89
80% fb 0.1 M K_2CO_3	10.36	0.591	0.495	0.91
70% fb 0.1 M K_2CO_3	10.17	0.568	0.472	0.87
60% fb 0.1 M K_2CO_3	10.03	0.567	0.471	0.86
50% fb 0.1 M K_2CO_3	9.85	0.571	0.474	0.87
40% fb 0.1 M K_2CO_3	9.68	0.541	0.445	0.82
30% fb 0.1 M K_2CO_3	9.47	0.527	0.431	0.79
50% fb 0.1 M NH_4Cl	9.26	0.476	0.380	0.70
20% fb 0.1 M K_2CO_3	9.20	0.454	0.358	0.66
40% fb 0.1 M NH_4Cl	9.10	0.407	0.311	0.57
30% fb 0.1 M NH_4Cl	8.86	0.362	0.265	0.49
10% fb 0.1 M K_2CO_3	8.78	0.314	0.218	0.40
20% fb 0.1 M NH_4Cl	8.50	0.273	0.177	0.32
90% fb 0.1M KH_2PO_4	7.54	0.200	0.103	0.19
80% fb 0.1M KH_2PO_4	7.25	0.188	0.092	0.17
80% fb 0.1M KH_2PO_4	7.25	0.183	0.087	0.16
70% fb 0.1M KH_2PO_4	7.04	0.170	0.073	0.13
60% fb 0.1M KH_2PO_4	6.83	0.156	0.059	0.11
50% fb 0.1M KH_2PO_4	6.61	0.143	0.047	0.09
40% fb 0.1 M KH_2PO_4	6.42	0.136	0.039	0.07
30% fb 0.1 M KH_2PO_4	6.21	0.122	0.025	0.05
20% fb 0.1 M KH_2PO_4	5.91	0.112	0.016	0.03
10% fb 0.1M KH_2PO_4	5.31	0.099	0.002	0.00
50 mM HCl	1.41	0.096[a]	0.000	0.00

[a] Value used for A_{237}^{RSH}. [b] Value used for $(A_{obs} - A_{237}^{RSH})_{max.}$

Fig. S46: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for H-Cys-Val-Phe-NH_2 (**10**) (0.2 mM) with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S2.10.1 Synthesis of H-Cys-Val-Phe-NH_2 (10)



As with the other tripeptides this compound was synthesised using Fmoc solid phase techniques. Rink amide MBHA resin (100-200 mesh, 0.436 mmol/g) (0.69 g, 0.30 mmol) was swollen for 1 h in DMF (5 mL). The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Phe-OH (0.29 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin. The resin was agitated for 1 h, drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Val-OH (0.25 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin. The

resin was agitated for 1 h, drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). Fmoc-Cys(mmt)-OH (0.46 g, 0.75 mmol), PyBOP (0.39 g, 0.75 mmol) and DIPEA (0.19 g, 1.50 mmol) were dissolved in DMF (3 mL) and added to the resin. The resin was agitated for 2 h drained and then rinsed with DMF (2 × 5 mL). The coupling was carried out a second time under the same conditions. The resin was then drained and treated with 20% (v/v) piperidine/DMF solution (5 mL, 2 × 10 min). The resin was subsequently drained and washed with DMF (3 × 5 mL). The resin was treated with Et_2O (5 mL, 30 min) drained and washed with Et_2O (2 × 5 mL). A solution of TFA:triisopropylsilane:water 95:2.5:2.5 (v/v) (3 mL) was added to the resin and agitated for 1 h. The solution was recovered and the resin washed with DCM (3 × 5 mL), the combined organic solutions were concentrated under reduced pressure. The peptide was triturated with ether (2 × 10 mL). The solvent was removed via freeze-drying to give the isolated peptide as a white solid (74 mg, 0.20 mmol, 67 % yield) ***m/z*** (ES+): 367 ($[M+H]^+$, 100%); **HRMS** (ES+): $[M+H]^+$ $C_{17}H_{27}N_4O_3S$ requires 367.1804, found 367.1799 (Fig. S47).

Fig. S47: High resolution mass spectrum of H-Cys-Val-Phe-NH_2 (**10**).

Elemental Composition Report **Page 1**

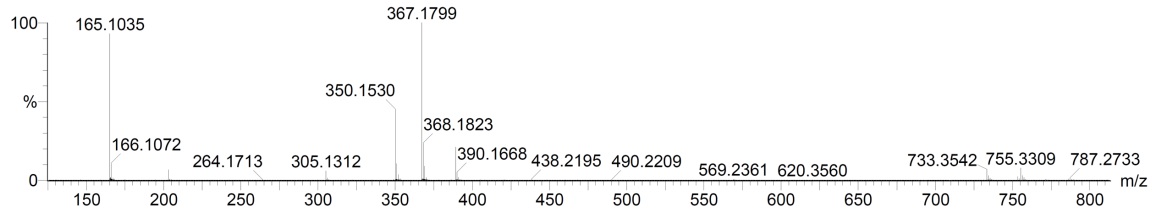
Single Mass Analysis
Tolerance = 3.0 mDa / DBE: min = -1.5, max = 100.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 5

Monoisotopic Mass, Even Electron Ions
922 formula(e) evaluated with 5 results within limits (up to 500 closest results for each mass)
Elements Used:
C: 0-60 H: 0-50 N: 0-8 O: 0-8 S: 0-2

QToF Premier

02-Oct-2019
WB5-145 216 (1.824) Cm (216:223)

1: TOF MS ES+
2.03e+005



Minimum:				-1.5							
Maximum:		3.0	5.0	100.0							
Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	i-FIT (Norm)	Formula				
367.1799	367.1804	-0.5	-1.4	6.5	1018.7	0.2	C17	H27	N4	O3	S
	367.1791	0.8	2.2	1.5	1020.3	1.8	C16	H31	O7	S	
	367.1810	-1.1	-3.0	15.5	1032.6	14.0	C25	H23	N2	O	
	367.1770	2.9	7.9	11.5	1031.6	13.0	C20	H23	N4	O3	
	367.1829	-3.0	-8.2	2.5	1032.1	13.6	C13	H27	N4	O8	

S3 Further Details of UV-Vis Spectrophotometric method

Benesch and Benesch have determined the dissociation constants K_a(A) – K_a(D) for cysteine and cysteine derivatives using a UV-Vis spectrophotometric method.[1] The dissociation constants are determined by monitoring how the absorbance from an n to σ* transition on the thiolate (λ_{max} ~ 237 nm) varied with pH. This method is based on a number of assumptions detailed below along with explanations for why each should not affect the results:

1. The molar extinction coefficients of thiolates in $^{+}H_3N$-R-S^{-} and H_2N-R-S^{-} are assumed to be identical. Benesch and Benesch reported that the molar extinction coefficient of *β*-mercaptoethylamine remained constant over the pH range 9.9 – 12.0.[S2] Within this pH range both the ammonium and thiol groups will transition from their conjugate acid and acid form to their base and conjugate base forms, respectively. Thus, the molar extinction coefficients of the thiolate in both species: ^{-}S-$CH_2CH_2NH_3^{+}$ and ^{-}S-$CH_2CH_2NH_2$, are equivalent as no discernable change was observed.
2. The molar extinction coefficients between the different cysteine derivative molecules examined are assumed to be approximately equivalent. This is imperative for the comparisons between speciation for the various cysteine methyl esters to be valid. This was experimentally verified to be the case (Section S4).
3. The observed absorbance at λ_{max} is solely due to the thiolate. This was again supported by studies of alanine by Benesch and Benesch. Negligible absorbances were observed at pH 7 when the amine of alanine is protonated and pH 12 when in base form.[S2] This gives confidence that any absorbance changes are due to changes at the thiolate and do not originate from any other moiety.
4. There are no competing reactions, such as oxidation or acid/base catalysed hydrolysis of ester / amide bonds occurring that will affect the absorbance changes. The impact of oxidation of thiolate to form disulfide bonds will be discussed below. Acid and base catalysed hydrolyses are not expected to be relevant because reaction half-lives for both of these reactions are significantly longer than the timescale (~ 2 min) of the UV-Vis spectrophotometric experiments:
    i. The second order rate constant for base catalysed hydrolysis of the ester ethyl acetate is k_{HO} = 0.087 M^{-1} s^{-1}, at pH 11.8 for an ester concentration of 0.2 mM, corresponds to a value of k_{obs} ~ 1.11 × 10^{-7} s^{-1} and a half-life of ~ 72 days.

ii. The second order rate constant for acid catalysed hydrolysis of the ester ethyl acetate is $k_{H+} = 1.07\ M^{-1}\ s^{-1}$ which at pH 1.5 and an ester concentration 0.2 mM corresponds to a value of $k_{obs} = 6.77 \times 10^{-6}\ s^{-1}$ and a half-life of ~ 28.5 h.

iii. At interim pH values, where acid and base catalysis are reduced, rate constants for hydrolysis will be further reduced.

S4 The Molar Extinction Coefficients of Cysteine Derivatives

One of the main assumptions behind this UV-Vis spectrophotometric approach to determining the acid dissociation constants is that the molar extinction coefficients between the different cysteine derivative molecules examined are approximately equivalent. This is imperative for the comparisons between speciation for the various cysteine methyl esters to be valid. To check whether this assumption was valid the molar extinction coefficients for cysteine **1**, cysteine methyl ester **2**, penicillamine methyl ester **3** and (4*S*)-mercaptoproline methyl ester **4** were determined.

Molar extinction coefficients are determined using the Beer Lambert Law:

$$A = \mathrm{c}.\varepsilon.l \qquad \text{(Eq. S2)}$$

where A is the absorption, c is the concentration of the compound in solution, ε is the molar extinction coefficient and l is the light path length the light travels through solution.

UV-Vis spectra were recorded in a 30 mM NaOH solution, in order to ensure that all thiol groups were in thiolate form (**iv**), and in the presence of the reducing agent TCEP (2.00 mM), in order to prevent oxidative disulfide bond formation.

S4.1 Determination of Molar Extinction Coefficient for Cysteine Methyl Ester (2)

A 0.1 M stock solution of cysteine methyl ester (**2**) was prepared from 1.72 mg (1.00×10^{-5} mol) of cysteine methyl ester.HCl and 100 μL of deionised water, a 0.100 M solution of *tris*(2-carboxyethyl)phosphine (TCEP) was prepared from 2.87 mg (1.00×10^{-4} mol) of TCEP and 100 μL of deionised water and a 10 mL solution of 30 mM NaOH at ionic strength I = 0.3 M (NaCl) was prepared from a 1.0 M volumetric NaOH standard solution.

As above, absorbance spectra between 200 – 300 nm were measured with a Cary 100 UV-Vis spectrophotometer thermostated at 25 °C. The background absorbance within the spectrophotometer was zeroed against air. 0.98 mL of the 30 mM NaOH solution was placed in 1.0 mL quartz cuvette and the absorbance spectrum of the solution was recorded. 20 μL of 0.1 M TCEP solution was added to give a 2.0 mM TCEP solution; the absorbance of the TCEP-containing solution was then determined. Different volume aliquots between 0.5 – 1.2 μL of

the 0.1 M stock solution of cysteine methyl ester were added to the cuvette to give 5.0×10^{-5} – 1.2×10^{-4} M solution of cysteine methyl ester. The absorption spectrum of cysteine methyl ester was recorded and the absorbance from the 2.0 mM TCEP solution was manually subtracted (Fig. S48).

The maximum absorbance (λ_{max} = 236) was plotted against concentration of cysteine methyl ester and the linear fit of the data to Eq. S2 gave the molar extinction coefficient (Fig. S49 and Table S15).

Fig. S48: UV-Vis spectra for 0.05 mM – 0.12 mM cysteine methyl ester (**2**) in 30 mM NaOH solution with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).

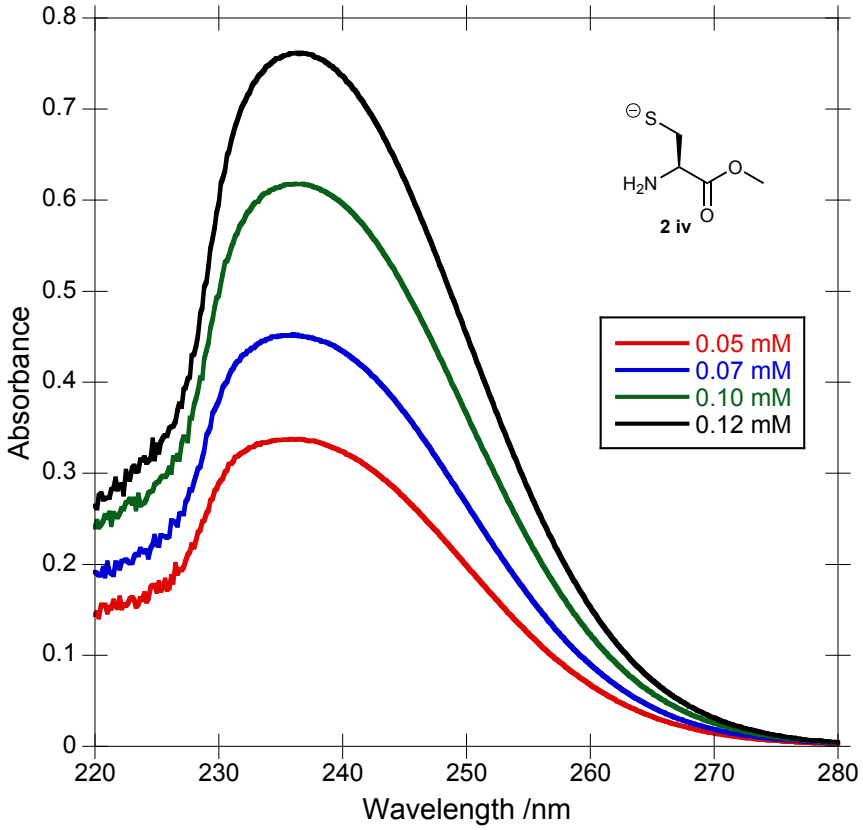
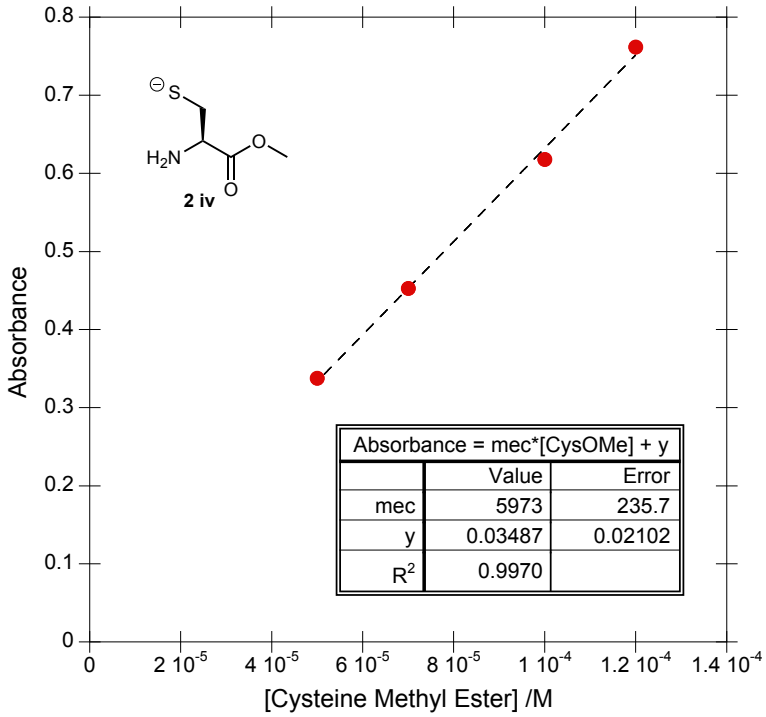


Table S15: The pH, λ_{max} and absorbance at λ_{max} values at 25 °C and ionic strength I = 0.3 M (NaCl) for different concentrations of cysteine methyl ester (**2**) used to determine the molar extinction coefficient at 25 °C and ionic strength I = 0.3 M (NaCl).

[Cysteine methyl ester] /mM	pH	λ_{max} /nm	Absorbance at λ_{max}
0.05	12.19	236	0.338
0.07	12.19	236	0.453
0.10	12.19	237	0.618
0.12	12.19	236	0.762

Fig. S49: Plot of maximum absorbance for 0.05 mM – 0.12 mM cysteine methyl ester (**2**) against concentration in 30 mM NaOH solution with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S4.2 Determination of Molar Extinction Coefficient for Penicillamine Methyl Ester (3)

A 0.1 M stock solution of penicillamine methyl ester (**3**) was prepared from 2.00 mg (1.00×10^{-5} mol) of penicillamine methyl ester.HCl and 100 µL of deionised water. The molar extinction coefficient was determined using the above method for cysteine methyl ester (**2**) (Fig. S50 + S51, Table S16).

Fig. S50: UV-Vis spectra for 0.05 mM – 0.15 mM penicillamine methyl ester (**3**) in 30 mM NaOH solution with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

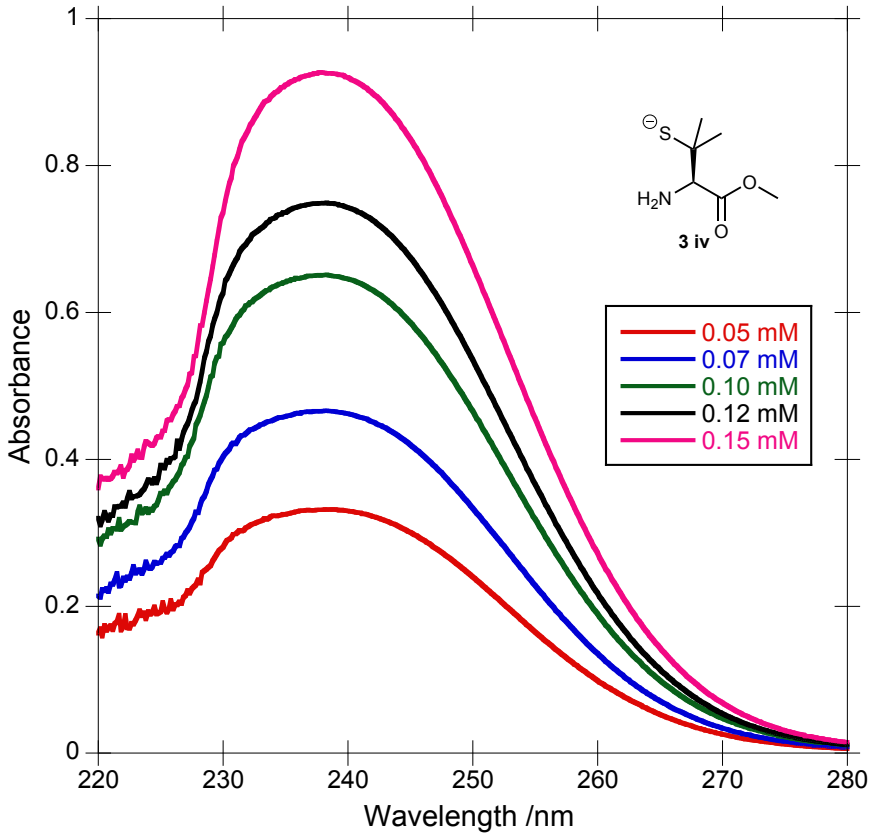
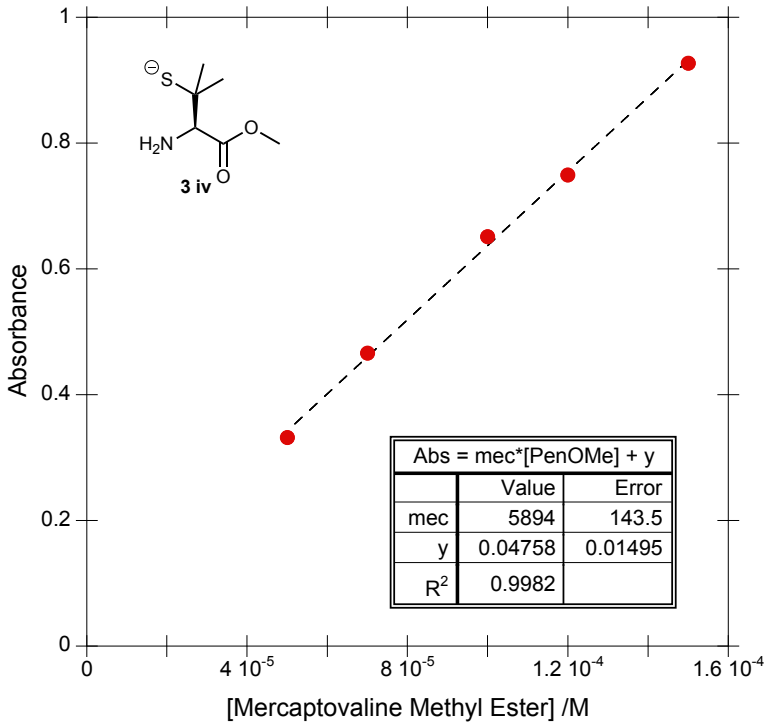


Table S16: pH, λ_{max} and absorbance values at 25 °C and ionic strength I = 0.3 M (NaCl) for different concentrations of penicillamine methyl ester (**3**) used to determine the molar extinction coefficient.

[Penicillamine methyl ester] /mM	pH	λ_{max} /nm	Absorbance
0.05	12.27	238	0.332
0.07	12.27	238	0.466
0.10	12.27	238	0.651
0.12	12.28	238	0.749
0.15	12.27	238	0.927

Fig. S51: Plot of maximum absorbance for 0.05 mM – 0.12 mM penicillamine methyl ester (**3**) against concentration in 30 mM NaOH solution with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S4.3 Determination of Molar Extinction Coefficient for the (4*S*)-Mercaptoproline Methyl Ester (4)

A 0.1 M stock solution of (4*S*)-mercaptoproline methyl ester (**4**) was prepared from 1.98 mg (1.00×10^{-5} mol) of (4*S*)-mercaptoproline methyl ester.HCl and 100 µL of deionised water. The molar extinction coefficient was determined using the above method for cysteine methyl ester (Fig. S52 + S53, Table S17).

Fig. S52: UV-Vis spectra for 0.05 mM – 0.12 mM 4*S*-mercaptoproline methyl ester (**4**) in 30 mM NaOH solution with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

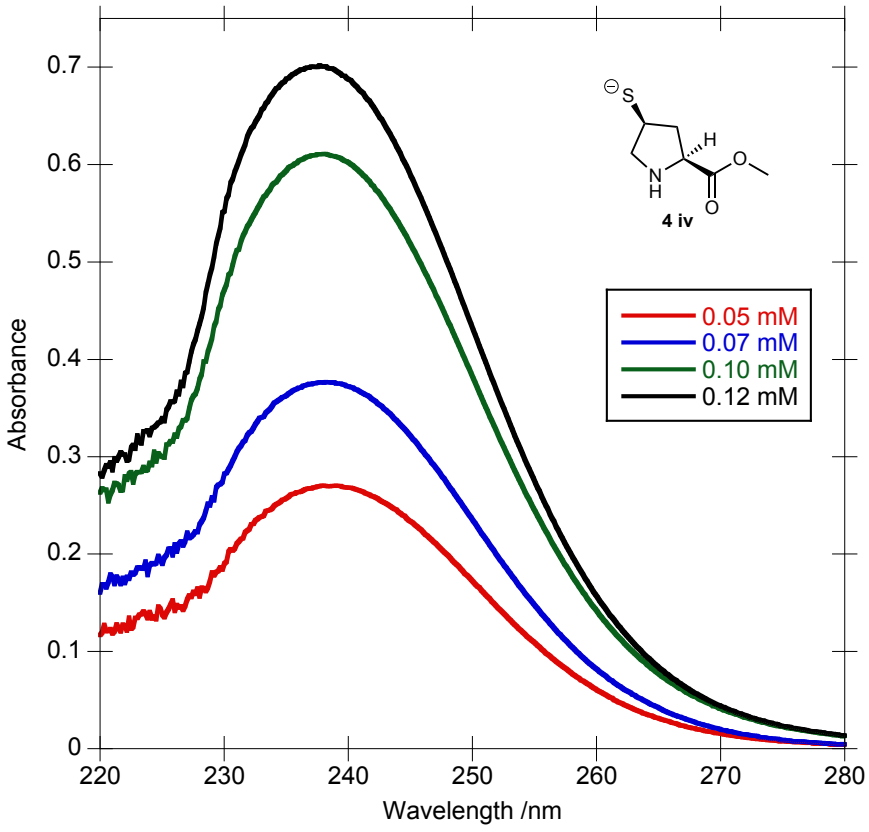
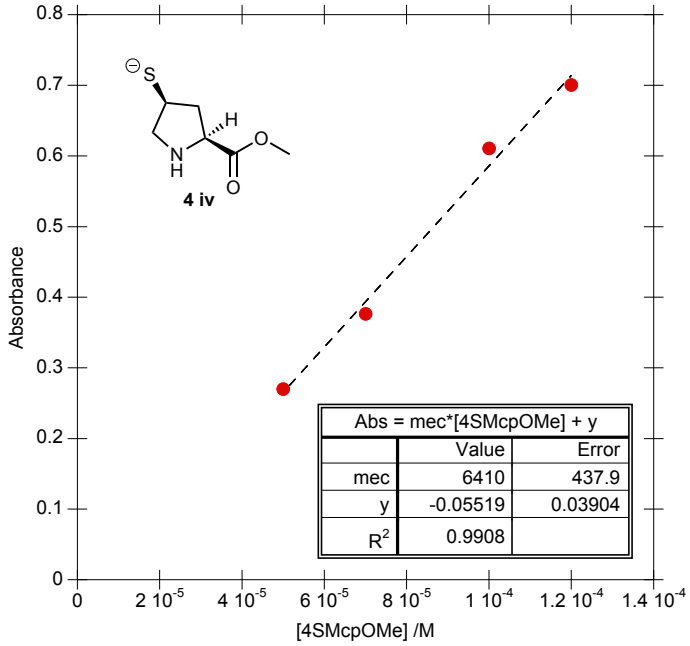


Table S17: pH, λ_{max} and absorbance values at 25 °C and ionic strength I = 0.3 M (NaCl) for different concentrations of 4*S*-mercaptoproline methyl ester (**4**) used to determine the molar extinction coefficient.

[4*S*-mercaptoproline methyl ester] /mM	pH	λ_{max} /nm	Absorbance
0.05	12.22	238	0.701
0.07	12.24	238	0.611
0.10	12.22	238	0.377
0.12	12.22	238	0.270

Fig. S53: Plot of maximum absorbance for 0.05 mM – 0.12 mM (4*S*)-mercaptoproline methyl ester (**4**) against concentration in 30 mM NaOH solution with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S4.4 Determination of Molar Extinction Coefficient for Cysteine (1)

A 0.1 M stock solution of cysteine (**1**) was prepared from 1.21 mg (1.00×10^{-5} mol) of cysteine and 100 µL of deionised water. The molar extinction coefficient was determined using the above method for cysteine methyl ester (**2**) (Fig. S54 + S55, Table S18).

Fig. S54: UV-Vis spectra for 0.05 mM – 0.12 mM cysteine (**1**) in 30 mM NaOH solution with 2.0 mM TCEP at 25 °C and ionic strength I = 0.3 M (NaCl).

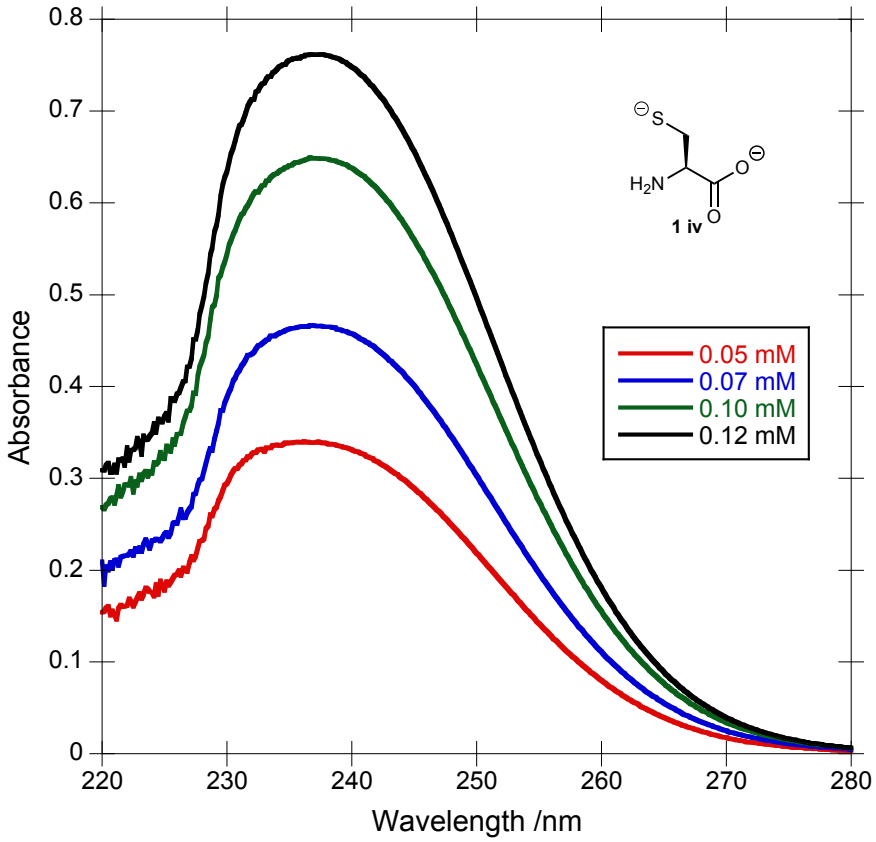
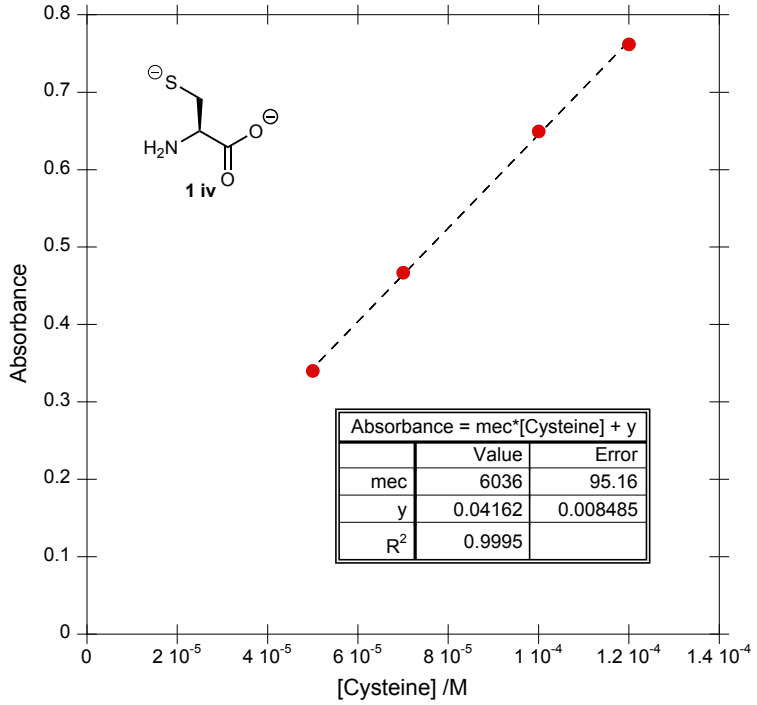


Table S18: pH, λ_{max} and absorbance values at 25 °C and ionic strength I = 0.3 M (NaCl) for different concentrations of cysteine (**1**) used to determine the molar extinction coefficient.

[Cysteine] /mM	pH	λ_{max} /nm	Absorbance
0.05	12.23	236	0.340
0.07	12.24	236	0.467
0.10	12.22	236	0.650
0.12	12.24	237	0.762

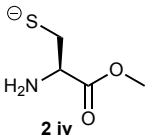
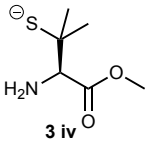
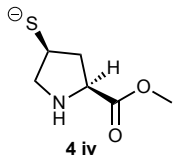
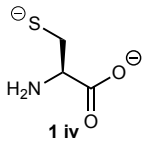
Fig. S55: Plot of maximum absorbance for 0.05 mM – 0.12 mM cysteine (**1**) against concentration in 30 mM NaOH solution with TCEP (2.0 mM) at 25 °C and ionic strength I = 0.3 M (NaCl).



S4.5 Comparison of molar extinction coefficients

The molar extinction coefficients, ε, for cysteine (**1**), cysteine methyl ester (**2**), penicillamine methyl ester (**3**) are closely similar (± 3 %). The ε value for (4*S*)-mercaptoproline methyl ester (**4**) has a < 9 % difference compared to the other values. The molar extinction coefficient determined for cysteine, ε = 6036 M^{-1} cm^{-1}, is in accordance with the literature value of ~ 6000 M^{-1} s^{-1}.[S2]

Table S19: Molar Extinction Coefficients for cysteine (**1**), cysteine methyl ester (**2**), penicillamine methyl ester (**3**) and (4*S*)-mercaptoproline methyl ester (**4**) at 25 °C and ionic strength I = 0.3 M (NaCl).

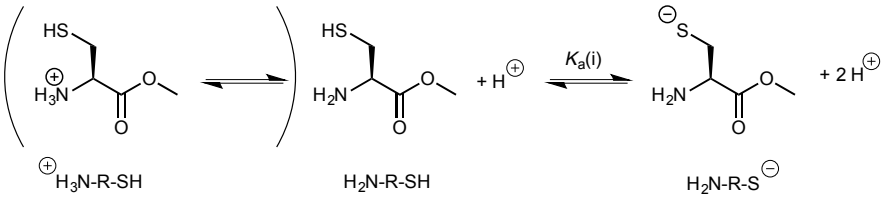
Compound	Molar Extinction Coefficient /M^{-1} cm^{-1}
 **2 iv**	5973
 **3 iv**	5894
 **4 iv**	6410
 **1 iv**	6036

S5 Fitting Outcome with an Alternative Speciation Model

A speciation model with four pK_as and four possible species in solution has been used to determine the pK of the thiol and ammonium (K_a(A) – K_a(D)) on the cysteine derivatives (Eq.2). Two alternative "Two pK_a + Three Species" models do exist where the pK_as of the thiol and ammonium are independent of each other (> 2 pK units apart).

The first of these "Two pK_a + Three Species" models has the ammonium pK_a below that of the thiol such that only a single thiolate species (H_2N-R-S^-) is present in solution (Scheme S1)

Scheme S1: "Two pK_a + Three Species" with the ammonium pK_a lower than that of the thiol.



Consequently, the K_a(A) of the thiol, K_a(i), may be found from Eq. S3:

$$K_a(\mathrm{i}) = \frac{[H_2N{-}R{-}S^-][H^+]}{[H_2N{-}R{-}SH]} \qquad \text{(Eq. S3)}$$

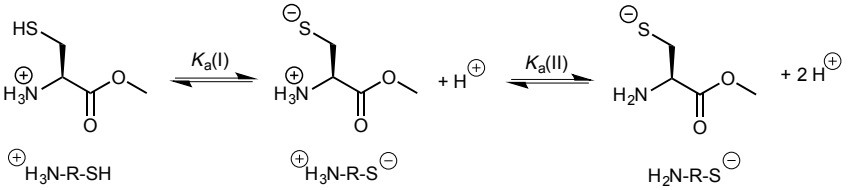
The fraction of thiol in thiolate form may therefore be found from Eq. S4:

$$f_{RS^-} = \frac{\frac{K_a(\mathrm{i})}{[H^+]}}{\frac{K_a(\mathrm{i})}{[H^+]} + 1} \qquad \text{(Eq. S4)}$$

The fit of the data for a penicillamine methyl ester (**3**) to Eq. S4 is shown in Fig. S56 (labeled as: 2 pK_a + 3 Species NH_3^+). The poor fit of the data to Eq. S4 will be discussed below.

The second "Two pK_a + Three Species" model has the thiol pK_a below that of the ammonium such that two thiolate species (^+H_3N-R-S^- and H_2N-R-S^-) may be present in solution (Scheme S2).

Scheme S2: "Two pK_a + Three Species" with the thiol pK_a (K_a(I)) lower than that of the ammonium (K_a(II)).



The concentration of both thiolate species at a given pH is therefore dependent upon the K_a(A) of the thiol (K_a(I)) (Eq. S5) and the K_a(A) of the ammonium (K_a(II)) (Eq. S6).

$$K_a(\mathrm{I}) = \frac{[\mathrm{H_3N^+{-}R{-}S^-}][\mathrm{H^+}]}{[\mathrm{H_3N^+{-}R{-}SH}]} \qquad \text{(Eq. S5)}$$

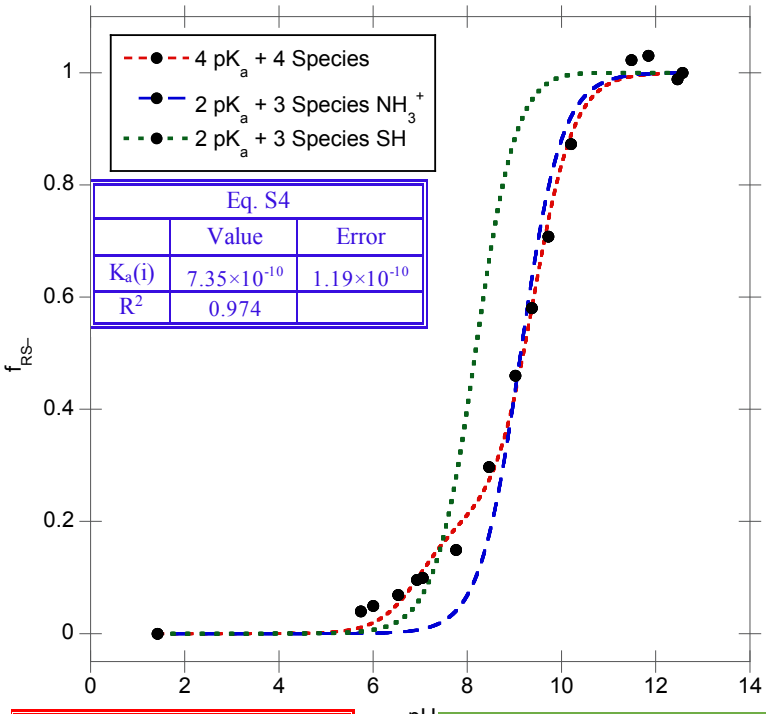
$$K_a(\mathrm{II}) = \frac{[\mathrm{H_2N{-}R{-}S^-}][\mathrm{H^+}]}{[\mathrm{H_3N^+{-}R{-}S^-}]} \qquad \text{(Eq. S6)}$$

The fraction of thiol in thiolate form may therefore be found from Eq. S7:

$$f_{\mathrm{RS^-}} = \frac{\frac{K_a(\mathrm{I})K_a(\mathrm{II})}{[\mathrm{H^+}]^2} + \frac{K_a(\mathrm{I})}{[\mathrm{H^+}]}}{\frac{K_a(\mathrm{I})K_a(\mathrm{II})}{[\mathrm{H^+}]^2} + \frac{K_a(\mathrm{I})}{[\mathrm{H^+}]} + 1} \qquad \text{(Eq. S7)}$$

The fit of the data for a penicillamine methyl ester (**3**) to Eq. S7 is shown in Fig. S56 (labeled as: 2 pK_a + 3 Species SH). The fit of the data to Eq. S7 will be discussed below.

Fig. S56: The variation of the fraction of thiol in thiolate form, f_{RS^-}, as a function of pH for penicillamine methyl ester (**3**) (0.2 mM) at 25 °C and ionic strength I = 0.3 M (NaCl). The fit of the data to the three possible pK_a and speciation models (Eq.2, Eq. S4 and Eq. S7) are shown.



Eq. S4		
	Value	Error
K_a(i)	7.35×10^{-10}	1.19×10^{-10}
R^2	0.974	

Eq. 2		
	Value	Error
K_a(A)	2.15×10^{-8}	7.33×10^{-9}
K_a(B)	8.52×10^{-8}	3.98×10^{-8}
K_a(D)	4.86×10^{-10}	4.49×10^{-11}
R^2	0.997	

Eq. S7		
	Value	Error
K_a(I)	6.66×10^{-9}	5.06×10^{-9}
K_a(II)	1.09×10^{-10}	1.62×10^{-9}
R^2	0.777	

The data for penicillamine methyl ester (**3**) was chosen for Fig. S56 because it demonstrated one of the poorest correlations to the "Four pK_a and Four Species" model, particularly in the pH region 5.9 – 7.9. Fig. S56 confirms that the "Four pK_a and Four Species" model is superior to both of the "Two pK_a + Three Species" models as the R^2 value for the fit is higher (0.997 vs. 0.974 and 0.777). The "Two pK_a + Three Species" model with a single thiolate species (Eq. S4, K_a(i)) has a better fit than that of the model with two thiolate species (Eq. S7, K_a(I) and K_a(II)). However, in the pH range 6.0 – 8.0 the fit of the "Two pK_a + Three Species" model with a single thiolate species is inferior to that of the "Four pK_a and Four Species" model. Attempts to fit the data of the other cysteine derivatives to either "Two pK_a + Three Species" models resulted in poor fits or in some cases failed to converge upon a solution.

S6 Variation in Species Concentration with pH

The percentage abundance for each of the four possible species $^{+}H_3N$-R-SH, $^{+}H_3N$-R-S^{-}, H_2N-R-SH and H_2N-R-S^{-} for the cysteine derivatives over the pH range 0 – 14 is shown in Fig. S57 – S66. Table S20 shows the percentage abundances for the various cysteine derivatives over a breadth of pHs at which native chemical ligations have been performed. The error in the values is ± 6 %.

Fig. S57: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for cysteine (**1**) at 25 °C and ionic strength I = 0.3 M (KCl).

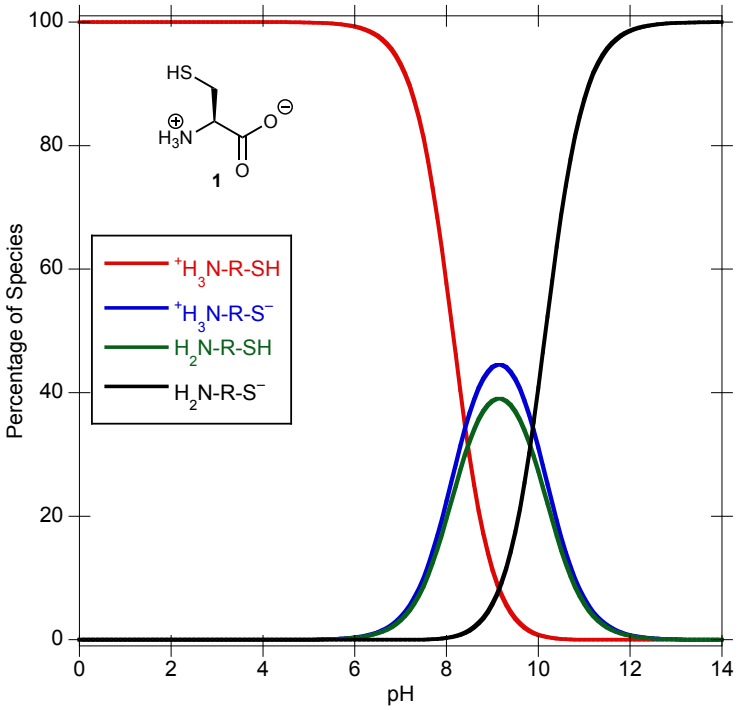


Fig. S58: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for cysteine methyl ester (**2**) at 25 °C and ionic strength I = 0.3 M (KCl).

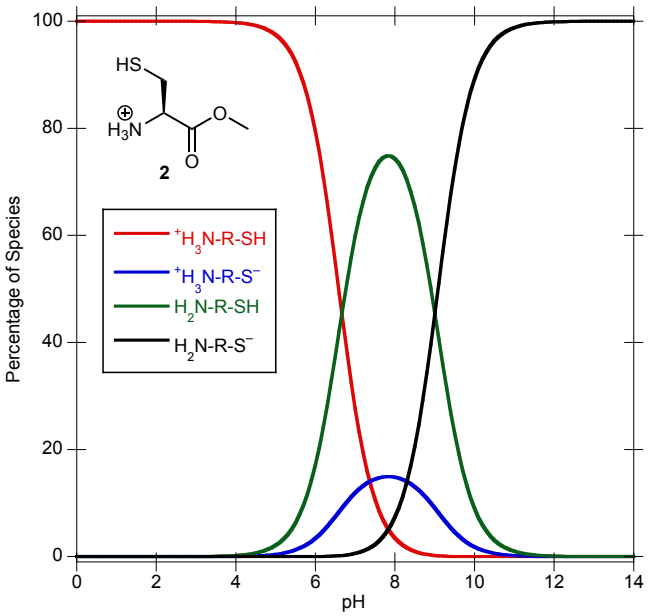


Fig. S59: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for penicillamine methyl ester (**3**) at 25 °C and ionic strength I = 0.3 M (KCl).

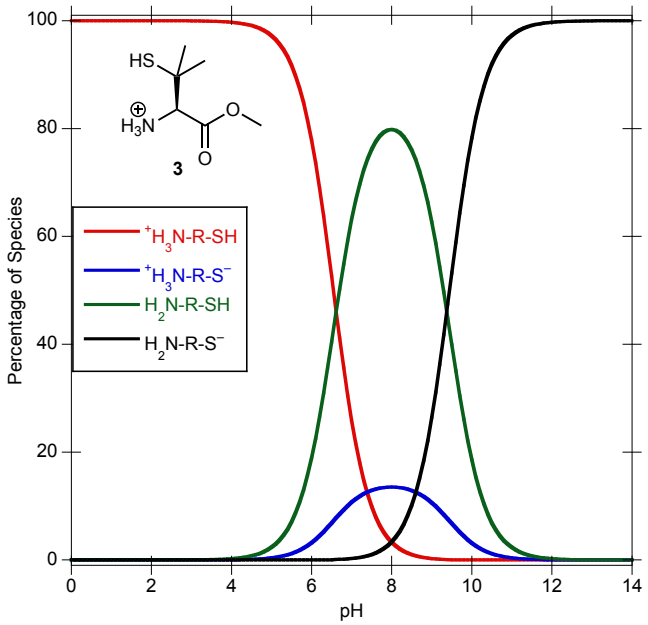


Fig. S60: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for (4*S*)-mercaptoproline methyl ester (**4**) at 25 °C and ionic strength I = 0.3 M (KCl).

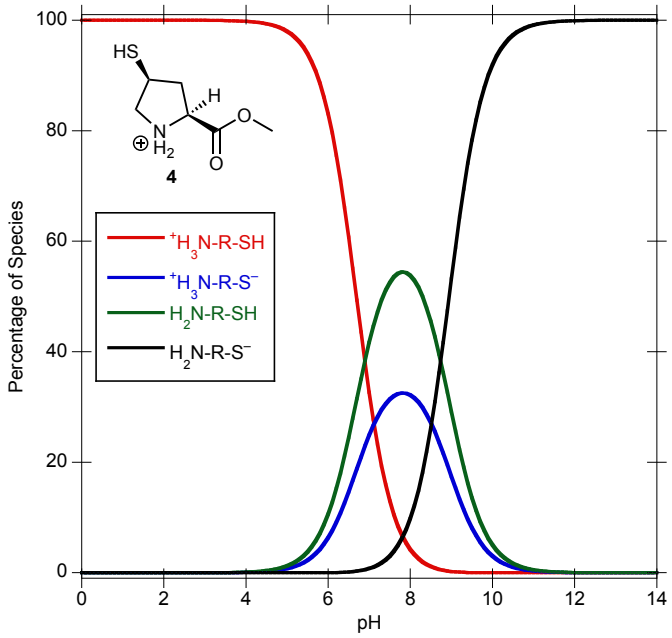


Fig. S61: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for H-Cys-Gly-OH (**5**) at 25 °C and ionic strength I = 0.3 M (KCl).

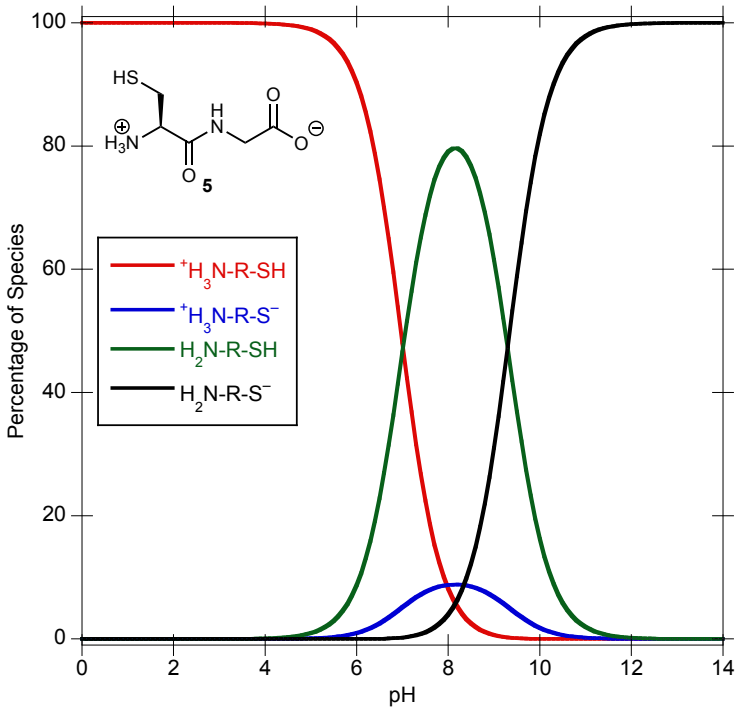


Fig. S62: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for H-Cys-Gly-Phe-NH_2 (**6**) at 25 °C and ionic strength I = 0.3 M (KCl).

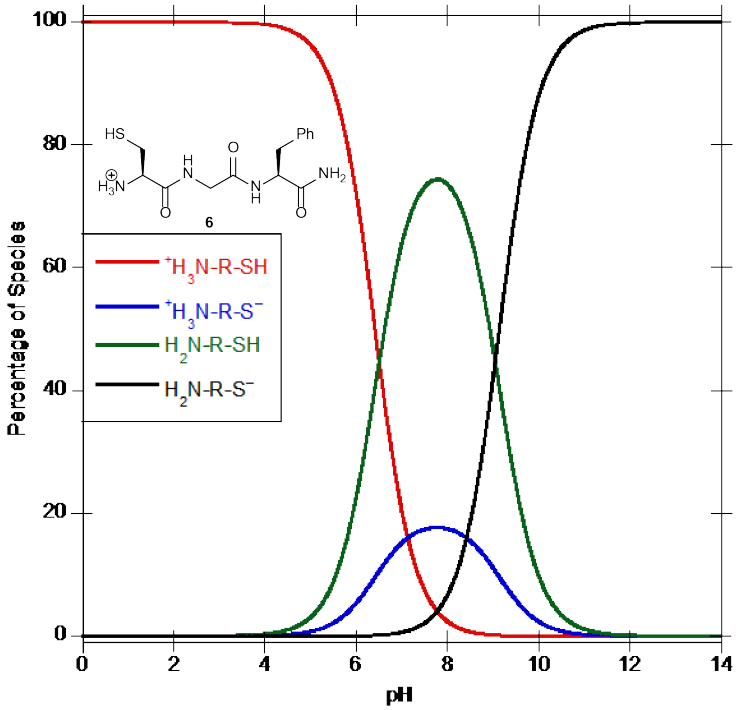


Fig. S63: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for H-Pen-Gly-Phe-NH_2 (**7**) at 25 °C and ionic strength *I* = 0.3 M (KCl).

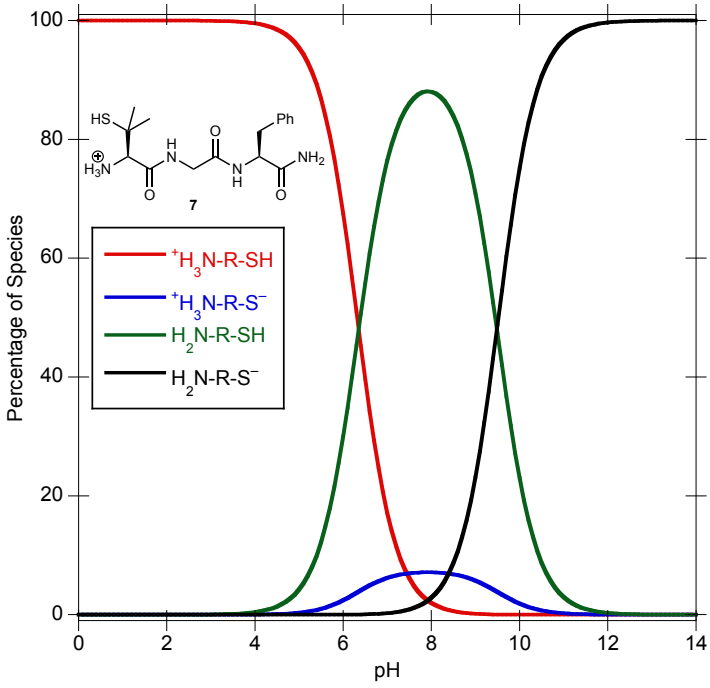


Fig. S64: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for H-(4S)-Mcp-Gly-Phe-NH_2 (**8**) at 25 °C and ionic strength I = 0.3 M (KCl).

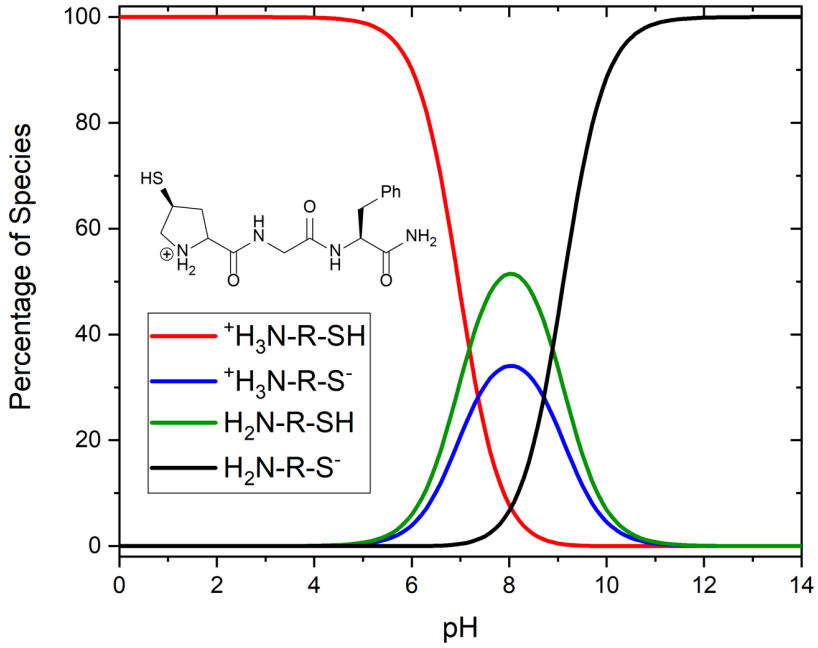


Fig. S65: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for H-Cys -Ser-Phe-NH_2 (**9**) at 25 °C and ionic strength I = 0.3 M (KCl).

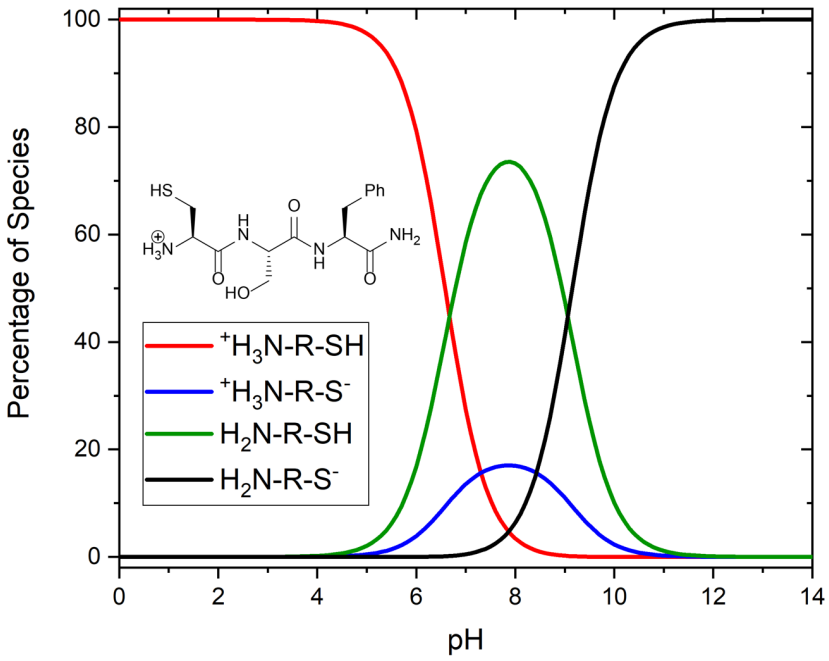


Fig. S66: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for H-Cys -Val-Phe-NH_2 (**10**) at 25 °C and ionic strength I = 0.3 M (KCl).

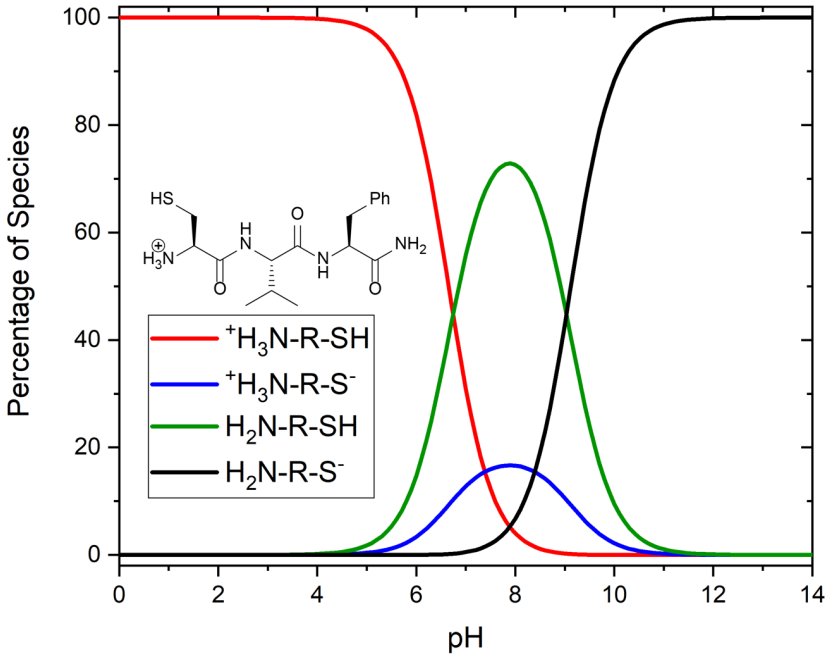
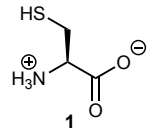
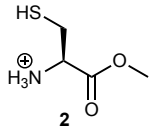
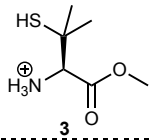
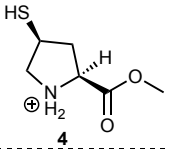
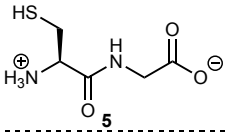
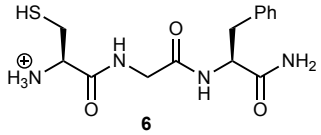
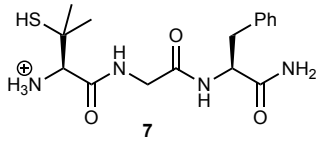
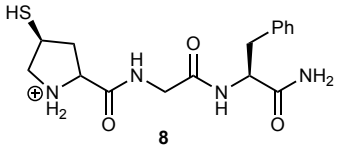
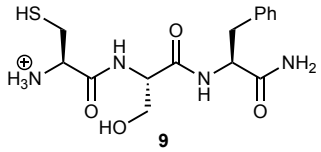
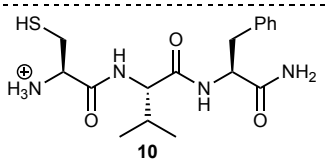


Table S20: The percentage abundance of each species for a range of representative pHs for cysteine derivatives studied in this work.[a] Data used to create pie charts in Main text Fig. 3.

Molecule	Species	pH 5	pH 6	pH 7	pH 8	pH 9
 1	$^{+}H_3N$-R-SH	100	99	93	58	11
	$^{+}H_3N$-R-S^{-}	0	0	4	22	44
	H_2N-R-SH	0	0	3	20	39
	H_2N-R-S^{-}	0	0	0	0	6
 2	$^{+}H_3N$-R-SH	97	78	26	3	0
	$^{+}H_3N$-R-S^{-}	0	4	13	15	9
	H_2N-R-SH	2	18	61	74	44
	H_2N-R-S^{-}	0	0	1	8	46
 3	$^{+}H_3N$-R-SH	97	78	26	3	0
	$^{+}H_3N$-R-S^{-}	0	3	11	14	11
	H_2N-R-SH	2	19	63	80	63
	H_2N-R-S^{-}	0	0	0	3	26
 4	$^{+}H_3N$-R-SH	98	83	33	4	0
	$^{+}H_3N$-R-S^{-}	1	6	25	32	17
	H_2N-R-SH	1	11	42	54	29
	H_2N-R-S^{-}	0	0	1	10	53
 5	$^{+}H_3N$-R-SH	99	90	48	8	1
	$^{+}H_3N$-R-S^{-}	0	1	5	9	7
	H_2N-R-SH	1	9	47	79	62
	H_2N-R-S^{-}	0	0	0	4	31
 6	$^{+}H_3N$-R-SH	96	72	20	2	0
	$^{+}H_3N$-R-S^{-}	1	5	15	17	11
	H_2N-R-SH	3	23	64	74	47
	H_2N-R-S^{-}	0	0	1	7	42
 7	$^{+}H_3N$-R-SH	95	68	17	2	0
	$^{+}H_3N$-R-S^{-}	0	2	6	7	6
	H_2N-R-SH	4	30	76	88	71
	H_2N-R-S^{-}	0	0	0	3	23
 8	$^{+}H_3N$-R-SH	99	88	42	6	0
	$^{+}H_3N$-R-S^{-}	1	5	22	32	17
	H_2N-R-SH	1	7	36	52	28
	H_2N-R-S^{-}	0	0	1	10	55
 9	$^{+}H_3N$-R-SH	97	79	28	3	0
	$^{+}H_3N$-R-S^{-}	0	4	14	17	11
	H_2N-R-SH	2	17	58	73	47
	H_2N-R-S^{-}	0	0	1	6	41
 10	$^{+}H_3N$-R-SH	98	82	31	4	0
	$^{+}H_3N$-R-S^{-}	0	3	13	17	11
	H_2N-R-SH	2	15	56	73	46
	H_2N-R-S^{-}	0	0	1	7	43

[a] The error in the values is ± 6 %.

S7 The effect of 6 M Guanidinium salt and 6 M Urea on the Absorbance in UV-Vis Experiments

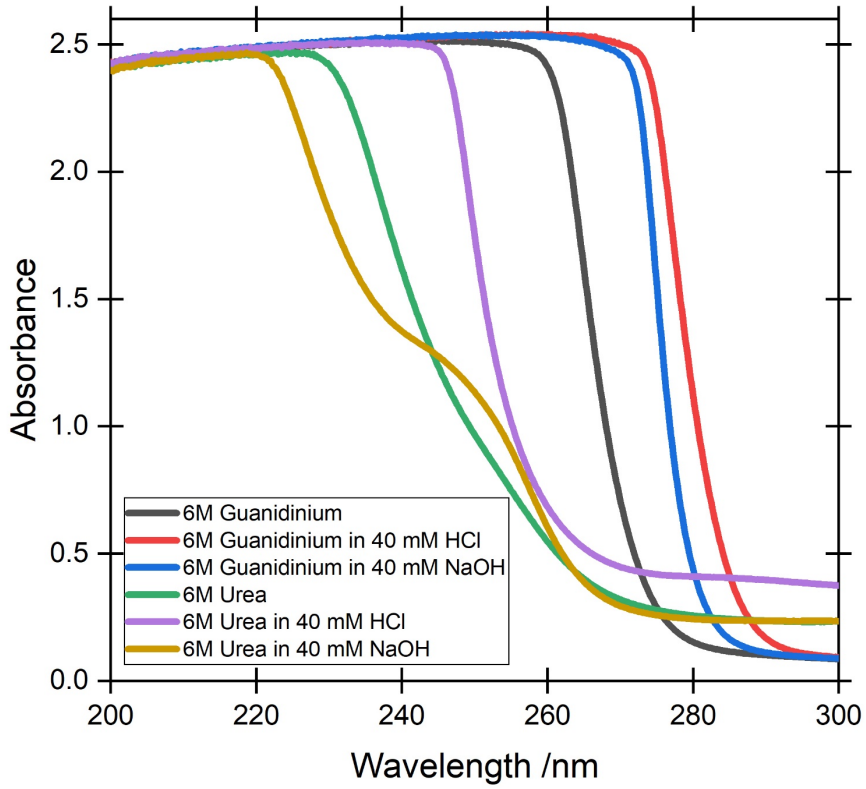


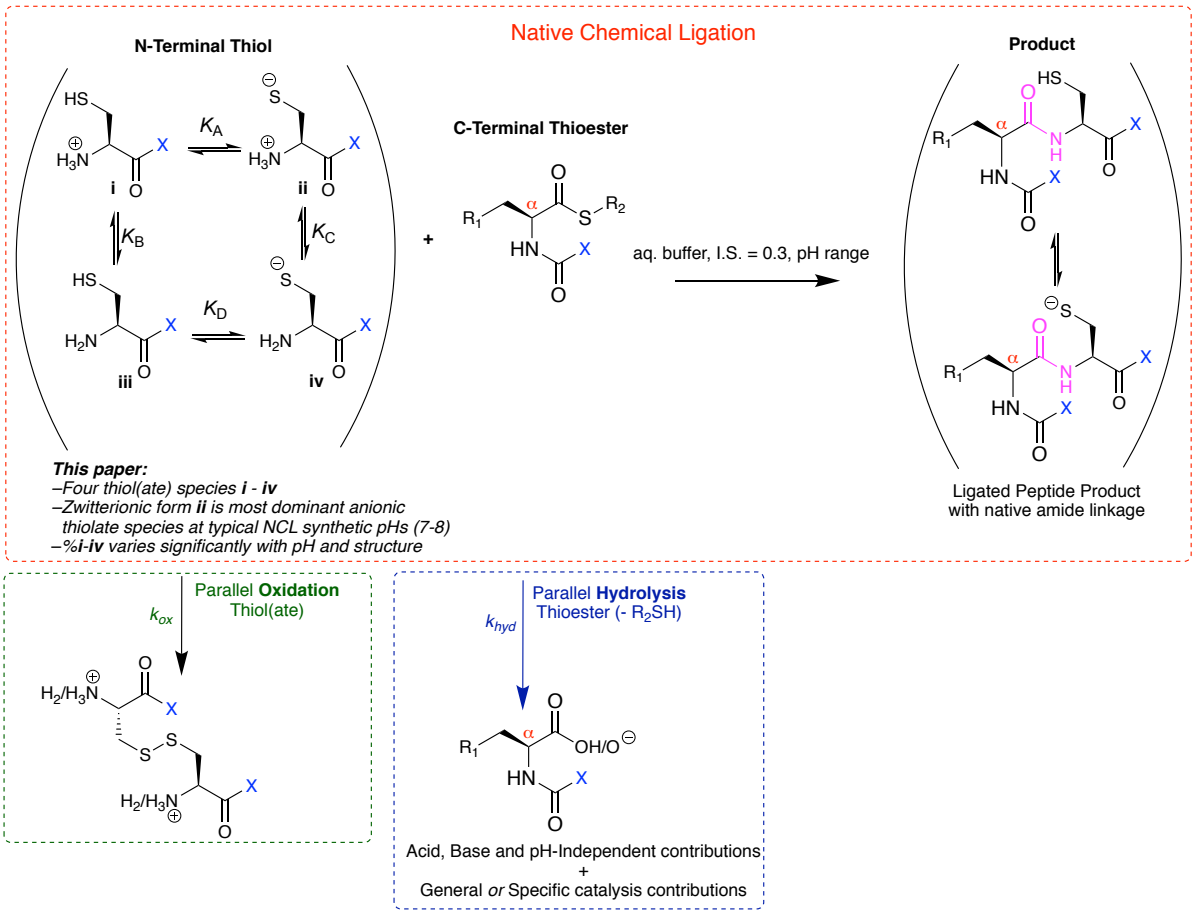
Fig. S67: UV-Vis spectra for 6 M Guanidinium salt and 6 M Urea in acidic, neutral and basic conditions at 25 °C.

In NCL a chaotropic agent can be added to suppress peptide folding/aggregation, which is important for longer peptide systems. Two common chaotropic agents used are 6 M Guanidinium HCl salt and 6 M Urea. We tested both to see whether we could also add these during our determination of the cysteine speciation by UV-Vis spectrophotometry. As Figure S67 shows, both additives either saturate or near-saturate absorbance at the wavelengths of interest (230 – 237 nm) and therefore were not included for our experiments.

S8 The Role of N-Terminal Speciation: Towards Correlation with NCL Rates

The overall NCL synthetic reaction requires N-terminal thiolated peptide and C-terminal peptide thioester substrates (Scheme S3).

Scheme S3: Synthetic NCL Reaction and Potential Competing Reactions



The observed rate of reaction of both substrates in aqueous solution under NCL conditions will have kinetic contributions from NCL (shown in red text in Eq. S8) and any parallel reactions, such as hydrolysis or oxidation,[S3-4] that contribute to substrate consumption.†

$$\text{Rate} = \left(\begin{array}{c} ([k_{NCL\text{-}i}\, f_i + k_{NCL\text{-}ii}\, f_{ii} + k_{NCL\text{-}iii}\, f_{iii} + k_{NCL\text{-}iv}\, f_{iv}] \times \textbf{[N-Thiol Peptide]} \times \textbf{[C-Terminal Thioester]}) \\ + \\ k_{hyd} \times \textbf{[C-Terminal Thioester]} \\ + \\ ([k_{ox\text{-}i}\, f_i + k_{ox\text{-}ii}\, f_{ii} + k_{ox\text{-}iii}\, f_{iii} + k_{ox\text{-}iv}\, f_{iv}] \times \textbf{[N-Thiol Peptide]} \end{array} \right) \quad \text{Eq. S8}$$

† Under the aqueous, buffered conditions of NCL, and for the relatively low peptide substrate concentrations used, a simple first order dependence on [N-Thiol Peptide] and [C-Terminal Thioester] is expected.

We have established herein that the NCL component of the overall rate will have contributions from four different thiol(ate) species **i-iv**. The fraction of each species, $f_{\text{i-iv}}$, can be calculated from the % **i-iv** and pK_a (A)– pK_a (D) values, and varies significantly with both pH and structural changes. To complete the full picture of the role of N-terminal thiol speciation in NCL, the different kinetic reactivities of **i-iv** ($k_{NCL\text{-}i}$ to $k_{NCL\text{-}iv}$) should also be determined towards typical C-terminal thioesters used in NCL. The combination of $f_{\text{i-iv}}$, $k_{NCL\text{-}i}$ to $k_{NCL\text{-}iv}$ and the quantification of any competing parallel reactions will allow full correlation of the abundance of species and reaction rates.

S9 References

S1. A. S. Hudson, 2015, *Efforts Towards the Total Synthesis of Labelled CCL2 Proteins and Dipeptide Chemotaxis Inhibitors*. Doctoral thesis, Durham University, available online at http://etheses.dur.ac.uk/11471/

S2. R. E. Benesch and R. Benesch, *J. Am. Chem. Soc.*, 1955, **77**, 5877.

S3. A. C. Conibear, E. E. Watson, R. J. Payne and C. F. W. Becker, *Chem. Soc. Rev.* 2018, **47**, 9046.

S4. P. J. Bracher, P. W. Snyder, B. R. Bohall and G. M. Whitesides, *Orig. Life Evol. Biosph.* 2011, 41, 399.